

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE INFORMÁTICA



EYECOLORS - IDENTIFICAÇÃO DA COR DOMINANTE EM FOTOGRAFIAS

Ruben Narciso Pavão

Mestrado em Engenharia Informática
Especialização em Sistemas de Informação

Dissertação orientada por:
Prof. Doutor Manuel João Caneira Monteiro da Fonseca

2016

Agradecimentos

A realização desta dissertação de mestrado contou com apoios importantes que passo a citar e agradecer.

Ao professor doutor Manuel João Caneira Monteiro da Fonseca pela orientação e paciência constantes.

Aos docentes que pelo seu método de ensino e pela sua personalidade me fizeram optar por esta área de especialização.

À faculdade de ciências de Lisboa e em particular ao Departamento de Informática e ao LaSIGE, por me facultarem os meios de aprendizagem e evolução.

Aos colegas durante o curso que se tornaram bons amigos, pela sua camaradagem, pelos momentos bons de trabalho e de melhor relaxe.

Aos meus amigos e colegas de casa que fizeram com que este percurso fosse cada vez mais prazeroso e menos desgastante.

À minha família, em especial aos meus pais, pela minha formação enquanto pessoa, pelo carinho, confiança, paciência e incentivos persistentes.

Aos meus braços por estarem sempre do meu lado, às minhas pernas por me apoiarem e aos meus dedos por eu poder sempre contar com eles.

Resumo

A pesquisa de imagens numa base de dados pode ser executada de diversas formas com base no seu conteúdo. Seja através de características como cores globais, cores locais, cores dominantes, texturas ou combinações de várias destas características. Os sistemas que utilizam as cores dominantes para as pesquisas de imagens fazem normalmente a sua identificação usando a perspetiva do sistema, e não tendo em conta a perceção que as pessoas têm das cores presentes numa imagem.

Neste trabalho temos como objetivo identificar as cores dominantes em imagens do ponto de vista dos utilizadores. Para tal, criamos um dataset de 100 imagens organizadas em 30 categorias e fizemos um estudo com utilizadores de modo a perceber que cores dominantes estes identificam nas imagens. Usamos ainda um eyetracker para determinar para que local na imagem as pessoas olham quando estão à procura das cores dominantes. Foram analisadas e agrupadas cores mencionadas pelos utilizadores, criados heatmaps, e fizemos algumas descobertas relativas à procura de cores dominantes através da perceção humana. Com base nestas informações desenvolvemos algoritmos para identificar as cores dominantes em imagens, do ponto de vista dos utilizadores. Um algoritmo simples que conta as cores dos pixéis na imagem, e um algoritmo de votos obtidos através de blocos nos quais a imagem foi dividida, que elegem cada as suas cores dominantes. A perceção humana é introduzida nos algoritmos com a adição de fatores aplicados às cores nas imagens, apurados através do estudo com utilizadores. É ainda proposto um modo de quantização das cores numa imagem para 12 cores de uma paleta utilizada em motores de busca populares. Com base numa análise aos algoritmos e aos fatores, identificamos a melhor solução usar o algoritmo com votos e com fator de cores puras a 10, que obtém uma taxa de acordo de 58,7% com os resultados retirados do estudo com utilizadores, e uma subida de 4,7% em relação ao algoritmo sem votos e sem fatores.

Palavras-chave: cores dominantes, imagens, perceção humana, estudo com utilizadores, eyetracker

Abstract

Image search in a database can be performed in different ways based on its content. Either through features like global colors, local colors, dominant colors, textures or a combination of several of these characteristics. Image search systems based on dominant colors often do the dominant colors identification through the system perspective, without considering the human perception of the colors in an image.

Our objective in this work is to identify the dominant colors in images from the human point of view. To that end, we created a dataset of 100 images organized in 30 categories and we made a study to see what dominant colors users identify on each image. We also used an eyetracker to determine the regions in the images that people look to, while searching for dominant colors. We analyzed and grouped colors mentioned by users, we created heatmaps and made some discoveries about dominant colors search through human perception. Based on this information, we developed algorithms to identify the dominant colors in images from the users' point of view. One is a simple algorithm that counts the colors on the image's pixels, the other is a voting algorithm where the votes are obtained from blocks, which were formed through an image division, and each block elects their dominant colors. The human perception is obtained through the addition of factors applied to the colors on the images, identified during the initial study with users. We also propose a way of making quantization of the colors in an image to a smaller palette of 12 colors often used in popular search engines. We tested the algorithms and factors and identified the voting algorithm with a 10 value for the factor of pure colors as the best solution. This configuration achieved a 58,7% accordance rate with the answers obtained on the users study, and a 4,7% growth from a simple algorithm without votes or factors.

Keywords: dominant colors, images, human perception, study with users, eyetracker

Conteúdo

Capítulo 1. Introdução.....	1
1.1. Motivação	1
1.2. Objectivos	1
1.3. Solução Desenvolvida	2
1.4. Contribuições.....	2
1.5. Estrutura do documento.....	2
Capítulo 2. Trabalho Relacionado.....	3
2.1. Percepção da Cor	3
2.2. Modelos de Cor	3
2.2.1. RGB	3
2.2.2. HSV.....	4
2.2.3. CIE L*a*b	4
2.3. Trabalhos sobre identificação de cor dominante	5
2.3.1. MPEG-7	5
2.3.2. Descritor baseado em cores dominantes e em pesos	6
2.3.3. Descritor baseado em cores dominantes e no contexto	7
2.3.4. Divisão da imagem para extração da cor dominante	8
2.3.5. Segmentação de imagem por cores principais	8
2.3.6. Análise e Discussão	9
2.4. Síntese.....	9
Capítulo 3. Estudo com Utilizadores e Eyetracker	11
3.1. Procedimento.	11
3.1.1. Conjunto de Imagens.....	11
3.1.2. Testes de Daltonismo	12
3.1.3. Realização dos Testes.....	12
3.2. Participantes.....	13
3.3. Paleta de Cores	14
3.3.1. Concordância de Cores.....	15
3.4. Regiões de Foco.....	16
3.4.1. Todas as Imagens	16
3.4.2. Categorias.....	17
3.4.3. Imagens Individuais	17
3.5. Discussão e Conclusões.....	18
3.6. Síntese.....	26
Capítulo 4. Algoritmo de identificação de cores segundo a percepção humana.....	27
4.1. Arquitetura.....	27
4.2. Pré-processamento.....	27
4.3. Quantização para 12 cores.....	28
4.4. Algoritmos	29
4.4.1. Algoritmo base	29
4.4.2. Algoritmo base com votos.....	30
4.4.3. Fatores para a percepção humana.....	30
4.4.4. Algoritmo de Percepção Humana	31
4.4.5. Algoritmo de Percepção Humana com votos.....	31
4.5. Síntese.....	31
Capítulo 5. Avaliação Experimental	33
5.1. Procedimento Experimental	33
5.2. Resultados dos Algoritmos Base	33

5.3.	Resultados de cada fator	33
5.3.1.	Análise do fator de Cores Quentes	34
5.3.2.	Análise do fator de Cores Puras	36
5.3.3.	Análise do fator de Centro da Imagem.....	38
5.3.4.	Análise do fator de Zonas Focadas	40
5.4.	Resultados dos algoritmos de percepção humana	42
5.5.	Análise e Discussão	44
5.6.	Síntese.....	44
Capítulo 6.	Conclusões e Trabalho Futuro	45
6.1.	Sumário da dissertação	45
6.2.	Trabalho Futuro	45
Bibliografia.....		47
Anexos.....		49

Lista de Figuras

Figura 2.1 – Modelo RGB.....	4
Figura 2.2 - Modelo HSV.....	4
Figura 2.3 - Modelo CIE L*a*b.....	5
Figura 2.4 - Para que serviam as normas MPEG-7.....	6
Figura 2.5 - Planificação do trabalho de Peng Wang.....	7
Figura 3.1 - Categorias das Imagens em http://press.liacs.nl/mirflickr/	11
Figura 3.2 - Exemplos de imagens incluídas no dataset	12
Figura 3.3 - Imagens do teste de Ishihara.....	12
Figura 3.4 - Cores mencionadas e mapeadas para as duas questões.....	14
Figura 4.1- Arquitetura da Solução.....	27
Figura 4.2- Espaço Cie L*a*b em duas dimensões.....	28
Figura 4.3 - Exemplo de imagens do conjunto submetidas à quantização para as 12 cores. ...	29
Figura 5.1 - Gráfico da análise do fator de cores quentes com cores dominantes.....	34
Figura 5.2 - Gráfico de análise do fator de cores quentes com cores de pesquisa.....	35
Figura 5.3 - Gráfico de análise do fator de cores puras com cores dominantes.....	36
Figura 5.4 - Gráfico de análise do fator de cores puras com cores de pesquisa.....	37
Figura 5.5 - Gráfico de análise do fator de cores no centro da imagem com cores dominantes.	38
Figura 5.6 - Gráfico de análise do fator de cores no centro da imagem com cores de pesquisa.	39
Figura 5.7 - Gráfico de análise do fator de cores em zonas focadas na imagem com cores dominantes.....	40
Figura 5.8 - Gráfico de análise do fator de cores em zonas focadas na imagem com cores de pesquisa.....	41

Lista de Tabelas

Tabela 3.1 - Tabela de Cores Identificadas nas Análises.	15
Tabela 3.2 - Tabela de Concordância para cores mais identificadas entre questões.	15
Tabela 3.3 - Heatmaps de todas as imagens para as duas análises.	16
Tabela 3.4 - Heatmaps de algumas categorias para as duas análises.	17
Tabela 3.5 - Heatmaps de algumas imagens para as duas análises.	17
Tabela 3.6 - Heatmaps foco em imagens com caras.	18
Tabela 3.7 - Heatmaps de natureza.	19
Tabela 3.8 - Heatmaps presença de luz.	20
Tabela 3.9 - Heatmaps de flores.	21
Tabela 3.10 - Heatmaps edifícios.	22
Tabela 3.11 - Heatmaps de letras presentes.	23
Tabela 3.12 - Heatmaps de cores puras.	24
Tabela 3.13 - Heatmaps de cores quentes.	25
Tabela 3.14 - Heatmaps de cores em zonas focadas.	26
Tabela 5.1 - Tabela com resultados de algoritmos base.	33
Tabela 5.2 - Tabela de análise ao fator de cores quentes com cores dominantes.	34
Tabela 5.3 - Tabela de análise ao fator de cores quentes com cores de pesquisa	35
Tabela 5.4 - Tabela de análise ao fator de cores puras com cores dominantes.	36
Tabela 5.5 - Tabela de análise ao fator de cores puras com cores de pesquisa.	37
Tabela 5.6 - Tabela de análise ao fator de cores no centro da imagem com cores dominantes.	38
Tabela 5.7 - Tabela de análise ao fator de cores no centro da imagem com cores de pesquisa.	39
Tabela 5.8 - Tabela de análise ao fator de cores em zonas focadas com cores dominantes. ...	40
Tabela 5.9 - Tabela de análise ao fator de cores em zonas focadas com cores de pesquisa. ...	41
Tabela 5.10 - Tabela de valores máximos para fatores com cores dominantes.	42
Tabela 5.11 - Tabela de valores máximos para fatores com cores de pesquisa.	42
Tabela 5.12 - Tabela de configurações de máximos a experimentar.	42
Tabela 5.13 - Configurações de máximos testadas no algoritmo sem voting.	43
Tabela 5.14 - Configurações de máximos testadas no algoritmo com voting.	44
Tabela 5.15 - Cores acertadas pelos algoritmos base.	44
Tabela 5.16 - Cores acertadas pelos algoritmos de percepção humana.	44

Lista de Anexos

Anexo 1 - Paleta Resultante do Estudo das Cores Dominantes.	49
Anexo 2 - Paleta Resultante do Estudo das Cores de Pesquisa.....	50
Anexo 3 - Guião para estudo com utilizadores.	51

Capítulo 1.

Introdução

A cor é uma das características visuais mais distintas, fazendo com que vários mecanismos ou sistemas para explorar, pesquisar e apresentar imagens aos utilizadores sejam baseados nas suas cores dominantes.

Neste trabalho propomos uma nova forma de extrair cores dominantes de imagens, que tem em conta o ponto de vista das pessoas. Desta forma, podemos usar as cores dominantes como descrição do conteúdo de imagens, facilitando, a sua pesquisa e exploração.

1.1. Motivação

Apesar de já existirem meios para pesquisar ou explorar imagens através das suas cores dominantes, estas normalmente são identificadas tendo em conta a perspetiva e os padrões do sistema e não a perceção humana das cores. Ou seja, tipicamente a cor dominante será aquela que ocupa a maior área da imagem, de entre o conjunto de cores (definição de cor dominante segundo a perspetiva do sistema). No entanto, do ponto de vista das pessoas as cores dominantes nem sempre são aquelas que têm mais pixéis.

1.2. Objectivos

Com este trabalho temos como objetivo principal o desenvolvimento de um algoritmo que identifique as cores dominantes numa imagem do ponto de vista da perceção humana. Para cumprir este objetivo, teremos de atingir os seguintes objetivos secundários:

- Perceber que cores dominantes as pessoas identificam nas imagens. Estas cores serão depois usadas para verificar se o algoritmo identifica as mesmas cores.
- Identificar os nomes das cores mencionadas pelas pessoas, para definir um conjunto reduzido de cores que poderão depois ser usadas para pesquisar imagens ou explorar coleções de imagens.
- Identificar as zonas das imagens para as quais as pessoas mais olham enquanto identificam as cores dominantes. Para isto usaremos um eyetracker (dispositivo que rastreia o movimento dos olhos).
- Relacionar estes dois últimos objetivos de modo a aferir quais as características das cores e das suas localizações, bem como os motivos que levaram as pessoas a olhar mais para elas.
- Avaliar o algoritmo usando os dados recolhidos no estudo inicial com utilizadores.

1.3. Solução Desenvolvida

Neste trabalho é proposto um algoritmo para identificar as cores dominantes segundo a percepção humana, que teve em conta um estudo prévio realizado com utilizadores, com o propósito de identificar as zonas mais visualizadas e cores mais mencionadas pelos utilizadores.

O algoritmo recebe uma imagem como entrada e devolve as suas cores dominantes, tendo em conta a percepção humana. Este algoritmo começa por fazer um pré-processamento para uniformizar o tamanho das imagens e reduzir o ruído existente nestas; faz uma quantização de cores para uma paleta de apenas 12 cores, ou seja converte cores para cores representantes de um vasto espetro possível destas; utiliza três modelos/espacos de cor para comparar as cores; e tem em conta fatores do estudo com utilizadores, dando assim mais peso às cores puras, cores quentes, cores presentes no centro da imagem, e cores presentes em zonas focadas na imagem.

1.4. Contribuições

Com o trabalho desenvolvido obtivemos as seguintes contribuições:

- Estudo com utilizadores para perceber para que regiões olham e que cores dizem perante um conjunto de imagens.
- Algoritmo de identificação de cores dominantes em imagens, com base na percepção humana, que utiliza descobertas feitas do ponto anterior.
- Algoritmo que faz a quantização de cores para uma paleta de 12 cores. A mesma paleta usada em motores de busca como Bing e Google.
- Conjunto de imagens, com as suas coordenadas recolhidas pelo eyetracker durante o estudo e os heatmaps resultante.

1.5. Estrutura do documento.

O restante documento está organizado da seguinte forma:

- Capítulo 2 – Neste capítulo explicamos sucintamente como é feita a percepção da cor pelos humanos e descrevemos alguns modelos de cor e trabalhos existentes para análise da cor dominante em imagens.
- Capítulo 3 – Neste capítulo descrevemos todo o estudo realizado com os utilizadores que ajudou a perceber como é que estes percebem as cores em imagens.
- Capítulo 4 – Apresentamos as nossas propostas para fazer a extração de cores dominantes de uma imagem, utilizando a percepção humana para chegar a esse fim.
- Capítulo 5 – Fazemos uma avaliação à precisão dos algoritmos propostos no documento e testes aos fatores aplicados nos algoritmos. Sugerimos o melhor algoritmo e configuração de fatores para determinar as cores dominantes nas imagens.
- Capítulo 6 – Apresenta as conduções e propostas de possíveis avanços que se podem realizar no algoritmo.

Capítulo 2.

Trabalho Relacionado

Neste capítulo apresentamos alguns trabalhos já existentes que lidam com cores e identificação da cor dominante em imagens. Para que estes sejam melhor interpretados e para que se possa perceber a sua importância, começaremos por descrever alguns modelos de cor existentes, considerados os mais adequados, e a forma como os humanos fazem a percepção da cor.

2.1. Percepção da Cor

Sendo a cor uma resposta sensorial provocada pela observação de estímulos produzidos por radiações de diferentes comprimentos de onda, a visão humana sente o espectro de cores usando combinações de informação presentes nas células localizadas nos olhos, são estas os cones e os bastonetes. As primeiras funcionam melhor com intensidades maiores de luz e são capazes de discernir cores, ao passo que os bastonetes são mais adaptados a situações de pouca luz e apenas detetam a intensidade desta.

Além da cor ser captada pelos olhos, a cor é uma representação interna a nível cerebral, e por isso a sua percepção depende fortemente de informação presente no cérebro, seja esta para fazer correção de cor mediante a iluminação presente, sobreposições de cores ou face a longas exposições de certas cores.

A percepção da cor é influenciada ainda pela cultura em que uma pessoa nasce, cresce e aprende, apesar de esta ser similar na maioria das culturas.

Como se pode ver, a percepção da cor pelo olho humano é um processo complexo e por isso é importante que os algoritmos de identificação de cores dominantes em imagens tenham isso em conta.

2.2. Modelos de Cor

Um modelo de cor é um modelo matemático abstrato, que descreve a forma como as cores podem ser representadas por tuplos de números. Os modelos de cor referem-se aos padrões de representação e estudo das cores, bem como às suas combinações, e servem ainda para estudar o comportamento das cores em diversos âmbitos. Existem vários modelos de cor [6] como por exemplo: RGB, CMYK, HSV, HSI, HSL, YUV, YIQ, YCbCr, CIE xyz, CIE $L^*u^*v^*$, CIE L^*a^*b .

Alguns destes modelos não conseguem representar todas as cores detetadas no espectro da visão humana, ao passo que outros representam mais cores do que um humano consegue perceber. Finalmente, existem alguns modelos que são perceptualmente uniformes, ou seja, a cor varia na mesma proporção das distâncias percorridas no modelo.

2.2.1. RGB

O RGB obtém o seu nome a partir das três cores primárias: Vermelho, Verde e Azul (“Red”, “Green” e “Blue”) que compõem o modelo aditivo, ou seja, onde a ausência de luz ou cor corresponde à cor preta e a mistura das três cores indica a presença da cor branca (como visto na Figura 2.1) As três cores variam entre os valores 0 e 1 quando representadas em valores decimais (ou entre 0 e 255, devido à conveniência de se guardar cada valor em 8 bits), onde o valor 1 representa a maior intensidade dessa

cor, e o valor 0 a ausência desta. Este modelo não representa todas as cores visíveis ao ser humano, não é perceptualmente uniforme e não facilita a escolha ou especificação de cores. É no entanto um bom modelo a ser usado para dispositivos que façam emissão ou detecção de luz.

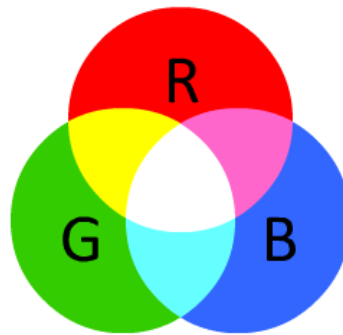


Figura 0.1 – Modelo RGB.

2.2.2. HSV

Este modelo tem como siglas a Matiz, Saturação e Valor (“Hue”, “Saturation”, e “Value”). A matiz, ou tonalidade, indica a cor pura (sem adição de branco ou preto) com saturação e iluminação máximas e permite fazer distinção das várias cores puras, com o seu intervalo entre os 0 e os 360 graus. A Saturação indica o grau de intensidade da cor e exprime-se num valor percentual, onde o valor 0% indica a inexistência de cor e o valor 100% indica uma cor saturada ou pura. O Valor traduz o brilho (luz emitida) de uma cor, também é expresso em valor percentual onde o valor 0% indica que a cor é muito escura ou preta e o valor 100% que a cor é saturada ou pura (ver Figura 2.2).

No HSV é mais fácil de escolher ou especificar uma cor do que no modelo RGB, por isso este modelo é mais usado em seletores de cores (Color Pickers) e por consequente em edição de imagem (Image Editing). Tem o problema de não ser perceptualmente uniforme, à semelhança do RGB.

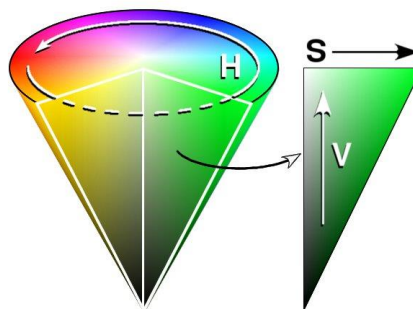


Figura 0.2 - Modelo HSV.

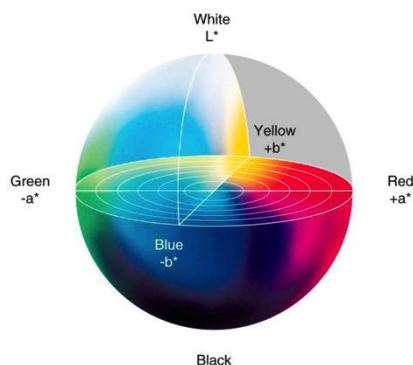
2.2.3. CIE L*a*b

O CIE L*a*b é um modelo com espaço de cor perceptualmente uniforme e de cores opostas, que retrata descobertas feitas nos anos 1960 que indicam que entre o nervo ótico e o cérebro os estímulos da cor da retina são convertidas entre distinções de claro e escuro, vermelho e verde, azul e amarelo. O CIE L*a*b, tem uma componente para a luz L, e duas para cores, a e b. Destas últimas, uma serve para definir se uma cor é mais verde ou vermelha e outra se a cor é mais amarela ou azul. Estes dois eixos

variam entre os 128 negativos e 127 positivos, ou entre 100 negativos e 100 positivos, dependendo da implementação e ocupação da cor branca no espaço de cor Lab.

Este modelo inclui mais cores do que as que são visíveis pelos humanos. Foi desenhado para se aproximar da visão humana, ou seja o seu gamut (gama de cores que pode ser processada) excede modelos como RGB e HSV. Foi criado ainda para que a cor não fosse afetada pelas propriedades do dispositivo, ao contrário dos modelos RGB, CMYK ou YCbCr.

O maior problema do modelo CIE $L^*a^*b^*$ é o facto de não ser fácil especificar ou escolher uma cor (ver Figura 2.3).



*Figura 0.3 - Modelo CIE $L^*a^*b^*$.*

2.3. Trabalhos sobre identificação de cor dominante

Com o crescimento da informação multimédia disponível em formato digital, foi necessário criar alguns sistemas que permitissem descrever o seu conteúdo automaticamente para que estes pudessem ser facilmente pesquisados e classificados. Para tal foram desenvolvidos vários tipos de descritores que estão encarregues de formar uma descrição do conteúdo presente num objeto multimédia, permitindo assim uma pesquisa rápida e eficaz. No caso das imagens, um dos descritores utilizados para descrever o seu conteúdo, e que é fácil de perceber pelas pessoas, é um composto pelas cores dominantes desta, DCD (Dominant Color Descriptor).

Nas secções seguintes, descrevemos alguns trabalhos cujo seu objetivo principal é a identificação de cores dominantes em imagens.

2.3.1. MPEG-7

MPEG-7 [3] [5] são um conjunto de normas e especificações, compostas pela ISO/IEC 15938, focadas na descrição de conteúdo multimédia, representando informação sobre o conteúdo e não o conteúdo em si ou a codificação deste, como era feito em normas anteriores a esta também do conjunto MPEG. O MPEG-7 faz uso de diferentes elementos como os “Descriptors” (D), “Description Schemes” (DSs) e “Description Definition Language” (DDL). Existem vários tipos de descritores para imagem neste conjunto de normas, sendo os descritores de cor dominante um deles. Estes permitem a especificação de um pequeno número de cores dominantes, bem como as suas propriedades estatísticas tais como a distribuição e a variância, com o objetivo de representar as cores presentes numa imagem ou região da imagem, de uma forma compacta e eficaz. Os “Description Schemes” servem para construir descrições mais complexas, especificando a estrutura e relações entre os descritores constituintes. A “Description Definition Language” permite uma definição flexível dos descritores e de “Description Schemes” feita em XML.

É importante mencionar que a extração e produção de um descritor não tinha forma mencionada nas normas do MPEG-7 quando este surgiu, nem tão pouco especificam um programa ou mecanismo de busca que faça uso da descrição, mas apenas a sua representação (como indica a representação na Figura 2.4).

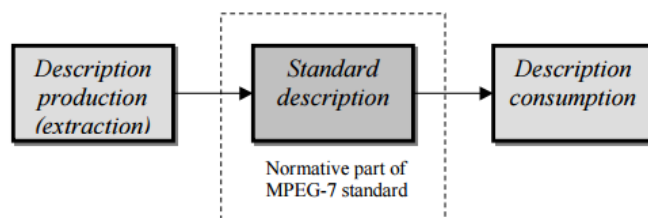


Figura 0.4 - Para que serviam as normas MPEG-7.

Em 2001 foram propostos alguns métodos de extração pelo MPEG-7, entre eles um focado no descritor baseado em cores dominantes, que extrai a distribuição das cores dominantes numa imagem. Este método é capaz de usar qualquer modelo de cor suportado pelo MPEG-7 (RGB, HSV, YCbCR e HMMD) e funciona agrupando as cores da imagem dentro de outro grupo de cores representativas em menor número (quantização). A extração é baseada no Algoritmo Generealizado de Lloyd (GLA), onde o centroide de cada grupo de cores é repetidamente encontrado depois de se agruparem as cores mais próximas a esse centroide (clustering). Este descritor consiste na percentagem, variância e coerência espacial das cores dominantes. A norma define um máximo de 8 cores dominantes, por imagem e usa a distância euclidiana para medir a similaridade entre os grupos de cores. De salientar que não existe um conjunto pré-definido de cores. As cores dominantes resultantes veem da junção das cores durante o clustering.

A extração no MPEG-7 não está adaptada a modelos que representam todas as cores percebidas por um humano, nem tem qualquer intervenção da percepção humana na sua conceção, problemas que o nosso trabalho visa em resolver. Tem no entanto o agrupamento de várias cores presentes numa imagem em pequenos grupos de cores (clusters), uma técnica que pode ser útil para o nosso trabalho, por exemplo para agrupar cores com a mesma matiz.

2.3.2. Descritor baseado em cores dominantes e em pesos

Em 2013 Talib [1] propôs um novo descritor baseado em cores dominantes para ser usado em pesquisa e exploração de imagens. O seu objetivo principal era obter uma descrição dos objetos presentes nas imagens através das suas cores dominantes. No entanto esta tarefa torna-se difícil quando a cor de fundo tem a maior percentagem entre as cores dominantes na imagem. Para resolver isto o autor partiu de algumas conclusões de estudos e consequentes suposições:

- Fotografos tendem a centrar um objeto numa fotografia.
- Imagens têm maioritariamente o objeto no seu centro e esse objeto ocupa, grande parte das vezes, 25% do espaço total da imagem.
- As cores de fundo aparecem principalmente nos cantos e nas bordas das imagens.

Assim, o algoritmo consegue distinguir cores dominantes para os objetos e para o fundo. Com base nesta informação são atribuídos pesos a cada cor dominante para que se consiga extrair informação das partes relevantes na imagem.

Esta forma de pensar inclui algumas vantagens mas também ideias pré-concebidas e costumes resultantes da ação humana que não são válidos (ex. elementos centrados nas fotografias). Além disto, existem alguns problemas nomeadamente o facto de um objeto poder tocar nas bordas da imagem, a cor de fundo ser da mesma cor que o objeto, ou uma linha de cor diferente da de fundo que contorne a imagem.

Enquanto os autores identificaram os locais da imagem relevantes para extrair cores dominantes a partir de suposições, no nosso trabalho, vamos identificar esses locais através do estudo com utilizadores, resultando assim numa forma mais próxima e fidedigna de como os utilizadores vêem as cores dominantes. Vamos, no entanto tentar aproveitar a ideia usada por Talib de considerar pesos diferentes para diferentes zonas da imagem, depois de termos feito o estudo com os utilizadores e apurado para que zonas estes olham mais.

2.3.3. Descritor baseado em cores dominantes e no contexto

Wang et al. [4] propuseram uma forma de extrair as cores dominantes através de um descritor, mais rico e eficiente, que tem em conta a informação contextual presente na imagem. Para isso, utilizaram paletes com 12 cores percecionadas e nomeadas por humanos. Conjuntos de tons para cada diferente cor foram submetidos a modelos de mistura Gaussiana (GMM), de modo a descobrir a probabilidade de um pixel pertencer a certa cor. Para uma dada imagem são criados 12 mapas de afiliação baseados nas probabilidades de um pixel pertencer a uma certa cor. Para cada mapa verifica-se se a probabilidade destes pixels pertencerem à cor é maior do que um threshold pré-definido. Para os pixels acima do threshold é analisada a região vizinha, o que resulta num mapa binário para cada cor dominante.

No último passo (ver Figura 2.5), o algoritmo, tem em conta que os humanos prestam mais atenção aos objetos salientes nas imagens. Para isso, usa um mapa de saliência para localizar os objetos nas imagens, e atribui uma bounding box ao objeto detetado. Depois de diferenciar as partes relevantes da imagem e o seu fundo são recalculadas as cores dominantes.

Este trabalho tem a vantagem de usar paletes RGB com várias cores percecionadas e diferenciadas por humanos no início do seu algoritmo. Partem ainda do pressuposto que os objetos marcam mais uma imagem, embora isso não esteja comprovado ou validado. A utilização da bounding box após aplicado o mapa de saliência nas imagens ajuda a apurar as cores perceptualmente dominantes.

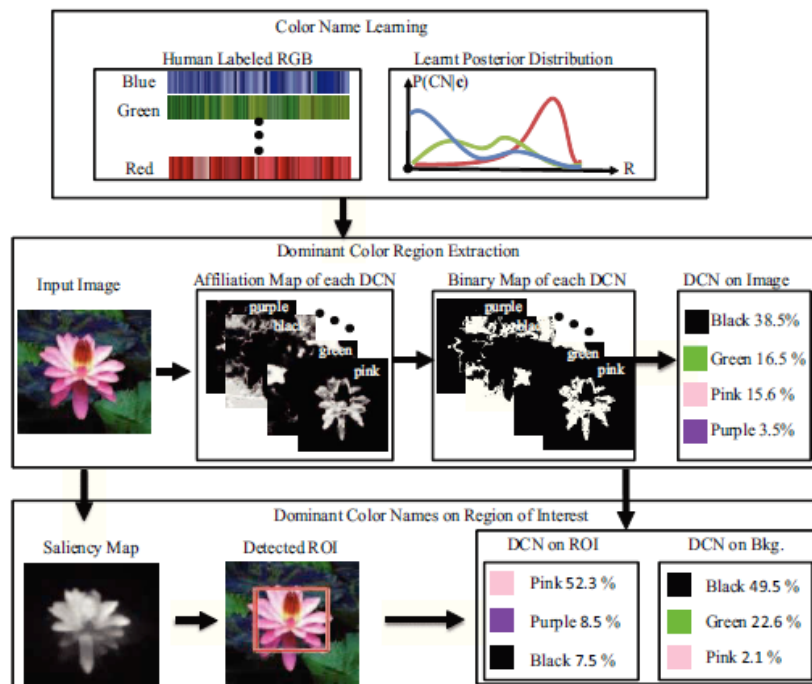


Figura 0.5 - Planificação do trabalho de Peng Wang.

2.3.4. Divisão da imagem para extração da cor dominante

Yang Lifang et al. [2] propuseram um algoritmo para detecção de similaridades entre imagens, que usa a distância entre as cores dominantes e as suas percentagens. Para tal combinam a existência da cor dominante com a distribuição da cor, dividindo a imagem numa matriz de 4x4. Para cada bloco da imagem é calculada a cor dominante do seguinte modo: para cada pixel num bloco é “quantizada” a sua cor e depois é calculado um histograma para o bloco de cores quantizadas para determinar a percentagem de cada cor no bloco. Cada bloco terá a sua cor dominante, ou seja a cor com maior percentagem. No final são comparadas as cores dominantes de cada bloco e as suas percentagens. Este algoritmo tem o problema de descartar algumas cores que podem ser dominantes na imagem total e não para o bloco em questão. Uma maneira de tornar o algoritmo mais robusto seria a divisão da imagem em mais sub-blocos e detetar e extrair mais do que uma cor dominante por bloco. O algoritmo não tem também qualquer auxílio da percepção humana na sua conceção.

A divisão da imagem em vários blocos poderá vir a ser útil para o nosso trabalho, no entanto a melhor forma de a aplicar deverá ser determinada após o nosso estudo com as pessoas.

2.3.5. Segmentação de imagem por cores principais

Ramella e di Baja [11] apresentaram um algoritmo de segmentação de imagem, que é um processo chave no reconhecimento de padrões e na visão computacional. O seu algoritmo é composto de duas fases: a primeira identifica as cores principais na imagem e constrói segmentações de pixéis com base em cores próximas dessas cores principais na imagem. Nesta primeira fase é construído um histograma para identificar os seus picos de cor (“componentes ligados de cor com máxima ocorrência local”). Os que têm uma ocorrência menor que um parâmetro pré-definido não são analisados. Este parâmetro é atribuído para diminuir o tempo computacional evitando a detecção de várias áreas pouco significantes na segmentação da imagem. Os autores reconhecem que podem estar a descartar alguns pixéis, não os atribuindo a nenhuma das segmentações, e por isso tomam esses pixéis em conta na fase seguinte. Há ainda um segundo parâmetro que é atribuído perante a aceitação do utilizador sobre a máxima dissemelhança entre cores aceites, o que poderá criar mais ou menos segmentações.

Todos os picos são tratados como detentores de cores dominantes, contudo alguns picos em que as cores não diferem muito têm as suas segmentações transformadas num único grupo, o que leva que o número final de grupos ou seja cores representantes seja em menor número do que os picos encontrados.

O segundo passo consiste em fundir algumas segmentações com base no seu tamanho, de forma a reduzir regiões na imagem segmentada, bem como pixéis não pertencentes a qualquer segmentação, resultantes do passo anterior. O tamanho das áreas dos pixéis é comparado a outro parâmetro de entrada, e caso sejam menores que esse tamanho estipulado esses pixéis irão pertencer à maior área adjacente.

A escolha dos parâmetros é fundamental neste trabalho para que não haja muitos pixéis descartados de qualquer segmentação e obter um resultado tolerável. Os autores conseguiram mostrar, com os parâmetros certos, uma extração das regiões salientes nas imagens muito próxima da percepção humana; reduzindo em muito o número de cores presentes na imagem original transformando-as em algumas representantes. O processo que usam para unir as segmentações no segundo passo deste trabalho causa uma redução no número de cores representantes.

Será preferível no nosso trabalho e em caso de existência de parâmetros que estes não sejam mostrados ao utilizador mas internos ao algoritmo e com base nos resultados da experiência com as pessoas. A forma de realizar as segmentações nas imagens também será determinada após o nosso estudo, tendo em conta agora que estas podem ser aglomeradas como vistas no trabalho destas autoras.

2.3.6. Análise e Discussão

Através dos trabalhos descritos anteriormente, estudamos várias técnicas e práticas para a extração da cor dominante em imagens. Uns usam quantização das cores para reduzir o número de cores distintas na imagem, escolhendo cores representantes durante o processo. Esta técnica melhora o tempo que demora a correr um algoritmo posterior de extração de cor dominante e a memória necessária para tal. Geralmente associado à quantização mas nem sempre alguns trabalhos fazem uso de clustering, para agrupar as cores por grupos. No MPEG-7 utilizam a distância euclidiana em conjunto com modelos de cores para medir a distância, ou similaridade, que duas cores têm entre si. Em alguns destes trabalhos também é usada a segmentação da imagem, para que o trabalho de extração de cor possa ser repartido e até atribuir importâncias relativas às zonas da imagem, que são comparadas no final dos algoritmos. Certos trabalhos fazem uso de suavizações gaussianas (Gaussian Blur) para tornar os limites das zonas de cor e transições de cor mais suaves através da redução de detalhes. No trabalho de Peng Wang et al. fazem uso de bounding boxes para apontar zonas de relevo na imagem em conjunto com mapas de saliência de modo a identificar a posição e contorno de elementos nas boxes.

No entanto, quase nenhum dos trabalhos estudados utiliza a percepção humana na sua base e conceção do algoritmo (com exceção de Peng Wang et al. que utilizam paletes de cores identificadas por humanos). Quando existe a utilização da “percepção humana” esta é feita de forma errónea por meio de escolha de parâmetros para obter melhores resultados num dado algoritmo, ou de orientação tomada através de suposições, sem estudo prévio com humanos para as comprovar.

Para o nosso trabalho é importante que o algoritmo a implementar seja baseado num modelo de cor que seja perceptualmente uniforme (por exemplo o CIELab) para facilitar a quantização das cores e a segmentação das imagens. Segmentações nas imagens poderão ser feitas caso o estudo a realizar assim o indique. Estas segmentações poderão aglomerar grupos de cores com a mesma matiz e terem importâncias (pesos) diferentes mediante diversos fatores a descobrir com o estudo. Eventualmente poderão ser usados mapas de saliência e bounding boxes de elementos identificados nas imagens. Finalmente e reiterando, será importante que o algoritmo a desenvolver não seja baseado em suposições, mas sim na realidade que será identificada no estudo com utilizadores e que em caso de existência de parâmetros estes não sejam uma escolha para o utilizador, mas internos ao algoritmo.

2.4. Síntese

Neste capítulo descrevemos como é feita a percepção da cor pelas pessoas, descrevemos alguns dos modelos de cor (RGB, HSV e CIE L^*a^*b), bem como as suas particularidades e diferenças. Apresentamos e analisamos alguns trabalhos mais relevantes para a identificação de cores dominantes em imagens. Algumas das técnicas usadas por estes são a quantização de cores usando clustering; a importância relativa das diferentes zonas na imagem; o uso de mapas de saliência para identificar regiões de interesse na imagem; a segmentação da imagem; e suavizações gaussianas para tornar as transições de cores mais suaves.

Capítulo 3.

Estudo com Utilizadores e Eyetracker

Neste capítulo descrevemos o estudo realizado com utilizadores para recolher informação sobre o modo como estes identificam as cores dominantes em imagens, quais as cores que utilizariam para encontrar uma dada imagem, e quais as zonas da imagem para onde olham durante a escolha das cores. A informação recolhida é útil para o desenvolvimento do nosso algoritmo de identificação das cores dominantes em imagens. A seguir descrevemos como criamos o conjunto de imagens a usar no estudo, o processo do estudo e a análise dos dados obtidos.

3.1. Procedimento.

O estudo realizado consistiu em mostrar 100 imagem, num monitor e uma de cada vez, a cada utilizador e pedir-lhes para fazerem uma de duas tarefas. A metade dos participantes pedimos que identificassem até três cores dominantes e à outra metade pedimos até três cores que utilizariam para pesquisar a imagem. As cores mencionadas foram registadas pela pessoa que conduziu os testes e as coordenadas da imagem para onde as pessoas olharam foram registadas usando um eyetracker.

3.1.1. Conjunto de Imagens

As imagens necessárias para a realização do estudo foram recolhidas do site flickr.com. Todas as imagens seleccionadas têm a licença Creative Common (permitem a sua utilização) e uma categoria (tag) associada. As categorias utilizadas para obter as imagens foram inspiradas nas trinta categorias utilizadas no dataset <http://press.liacs.nl/mirflickr/> (ver Figura 3.1). Apesar de já existir este conjunto de imagens, não o usamos porque ele apenas contém imagens com tamanhos reduzidos (ex: 500x500), o que resultaria numa extração de informação pouco precisa sobre as zonas da imagem para onde as pessoas olham, ao colocar as pessoas em frente ao eyetracker.

Tag	# Images	Tag	# Images
sky	845	people	330
water	641	city/urban	308/247
portrait	623	sea	301
night	621	sun	290
nature	596	girl	262
sunset	585	snow	256
clouds	558	food	225
flower/flowers	510/351	bird	218
beach	407	sign	214
landscape	385	car	212
street	383	lake	199
dog	372	building	188
architecture	354	river	175
graffiti/ <u>streetart</u>	335/184	baby	167
tree/trees	331/245	animal	164

Figura 0.1 - Categorias das Imagens em <http://press.liacs.nl/mirflickr/>

As nossas imagens foram escolhidas executando uma busca avançada no Flickr usando as categorias listadas na imagem anterior, com orientação horizontal ou vertical, tamanho mínimo grande

e com modificação e uso comercial permitidos. Escolhemos quatro imagens por cada busca da seguinte forma: a primeira, terceira, quinta e sétima imagens. Nos casos em que algumas das imagens contivessem algum tipo de assinatura ou se notasse alguma modificação digital, estas eram substituídas por uma imagem válida, ainda não escolhida, numa posição imediatamente a seguir. No final desta recolha ficamos com um total de 120 imagens, das quais descartamos 20 para ficarmos apenas com 100. As imagens descartadas foram aquelas que se assemelhavam a outras já escolhidas em termos de contexto e cor. As imagens foram todas redimensionadas de modo a ficarem com a largura ou altura (maior dimensão destas duas) iguais à mesma dimensão do ecrã em que foram corridos os testes, redimensionando a outra dimensão proporcionalmente. Desta forma conseguimos uma conversão direta das coordenadas do ecrã (imagem) para as coordenadas do eyetracker. Na Figura 3.2 estão exemplos de imagens presentes no nosso dataset.

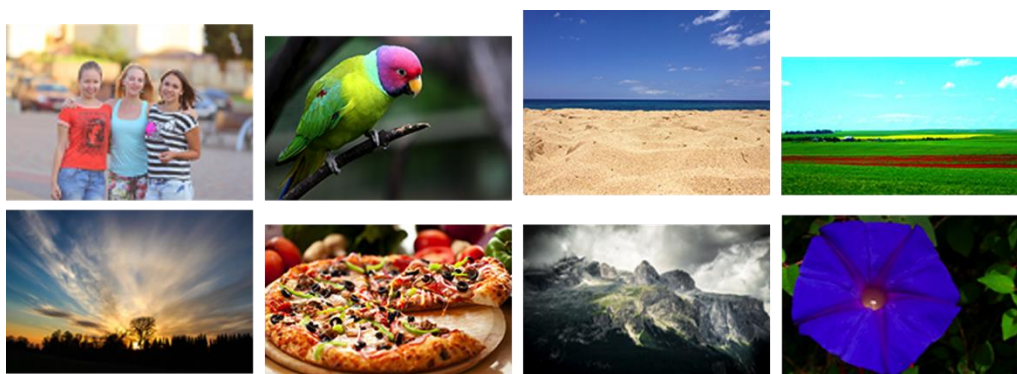


Figura 0.2 - Exemplos de imagens incluídas no dataset

3.1.2. Testes de Daltonismo

No início dos testes e para evitar a recolha de informação não útil para o estudo, mostramos três imagens do teste de Ishihara [7], para detetar a existência de daltonismo nos utilizadores. Este grupo de três imagens do teste de Ishihara (ver Figura 3.3) foi o mesmo para todos os utilizadores.

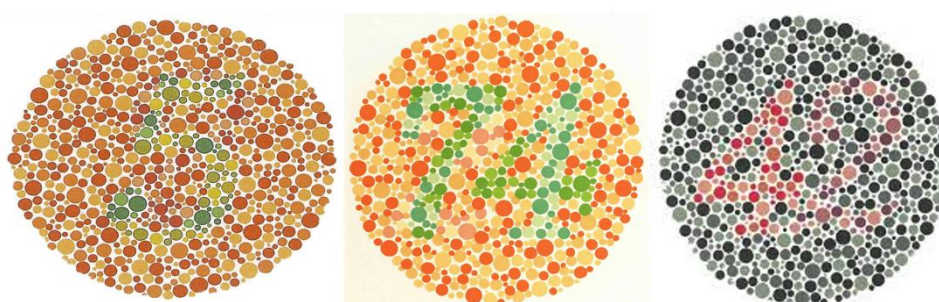


Figura 0.3 - Imagens do teste de Ishihara.

3.1.3. Realização dos Testes

Durante a realização dos testes, os participantes foram colocados em frente a um computador munido de um eyetracker para seguir e registar (para posterior análise) os movimentos dos seus olhos e as zonas das imagens para as quais estiveram mais tempo a olhar. Este eyetracker pertence à organização

TheEyeTribe¹, e consiste num modelo de captura que além de barato é fácil de transportar, instalar e oferece uma boa precisão para efeitos de estudos científicos. Para recolher os dados do eyetracker usamos as API's fornecidas pela TheEyeTribe, colocamos os participantes a uma distância compreendida entre 50 a 70 cms do ecrã e do eyetracker (que se encontrava exatamente por baixo do ecrã). A frequência de recolha de coordenadas do eyetracker foi ajustada para a maior possível (60Hz) o que resulta num maior número de coordenadas registadas para cada imagem durante o tempo em que esta esteve visível.

Cada utilizador foi exposto ao conjunto de cem imagens (dataset mencionado previamente nesta secção), sendo cada imagem mostrada durante sete segundos. Para cada imagem pedimos aos utilizadores para responderem a uma das duas perguntas:

Q1: “Quais as (até) três cores que identifica como dominantes nesta imagem?”

Q2: “Quais as (até) três cores que utilizaria para pesquisar esta imagem?”

Estas perguntas foram feitas para verificar se existia diferença nas respostas dos utilizadores e na essência da pergunta em si, uma vez que os resultados do estudo se destinam à criação de um algoritmo que faz a identificação de cores dominantes em imagens, que por sua vez facilita a pesquisa de imagens numa base de dados. Para evitar influenciar as respostas dos utilizadores não apresentamos uma listagem de cores, pedindo aos utilizadores que indicassem as cores livremente. Para diminuir as distrações e os desvios de olhares do ecrã os utilizadores diziam as cores verbalmente e a pessoa que conduzia os testes fazia o seu registo. As cores registadas para cada imagem foram depois agrupadas para cada questão. Este processo teve em conta o número de vezes que uma cor foi proferida em relação a uma imagem, o número de vezes que a cor foi proferida para as duas perguntas e a proximidade entre as cores indicadas.

No final do estudo foram comparadas as cores mencionadas com as coordenadas para as quais as pessoas mais olharam e com essa informação criamos paletes e heatmaps para podermos retirar mais conclusões, como veremos nas secções seguintes.

3.2. Participantes

Antes de começarmos o estudo e para cada utilizador que passou no teste de Ishihara recolhemos a sua informação demográfica, nomeadamente a idade, sexo e a utilização ou não de lentes de contacto ou de óculos (em alguns casos de estudo notou-se que alguns tipos de óculos causavam pequenos erros nas coordenadas registadas, que não teve muito efeito devido ao elevado número de participantes).

Para os participantes válidos fizemos uma calibração individual com o eyetracker, onde tinham que focar alternadamente entre doze pontos no ecrã através de um acompanhamento visual do percurso de uma circunferência entre esses pontos.

O estudo foi realizado com 40 utilizadores, divididos em dois grupos de 20 elementos cada um (grupo A e grupo B). O grupo A respondeu à questão Q1 e era composto por 14 utilizadores do sexo masculino e seis do sexo feminino, seis destas pessoas usavam óculos e apenas uma lentes de contato. Este grupo de pessoas tinha uma média de idades de 22 anos, e desvio padrão 2,86. O grupo B respondeu à questão Q2 e era composto por 12 pessoas do sexo masculino e oito do sexo feminino, sete usavam óculos e duas lentes de contato. A média de idades deste grupo era de 21 anos e desvio padrão de 2,96.

O teste foi publicitado entre alunos dos cursos de Licenciatura em Engenharia Informática e Licenciatura em Tecnologias Informação da Faculdade de Ciências da Universidade de Lisboa, bem como a conhecidos do autor deste trabalho.

¹ <https://theeyetribe.com/>

3.3. Paleta de Cores

Nesta secção diferenciamos e comparamos os dois tipos de análise de cores, dominantes (Q1) e de pesquisa (Q2), e mostramos como é que as cores foram agrupadas e os heatmaps foram criados.

Depois da recolha de coordenadas e cores, verificamos as cores proferidas para as duas perguntas. Com o objetivo de criar paletes pequenas que representassem as cores mencionadas para cada uma das questões, foi necessário reduzir o conjunto de nomes para as cores, mapeando certas cores ditas pelos participantes para outras mais representativas.

Para tal, experimentamos, e com sucesso, a palette sugerida por Ware [10], que utiliza doze cores (Preto, Branco, Cinzento, Vermelho, Amarelo, Laranja, Azul, Rosa, Roxo, Castanho, Verde e Cíano). Esta palette é baseada na palette JNS (“Just Not the Same”) identificada por Chang et al. [9], com a adição da cor cíano. As cores JNS representam as cores que conseguem ser nomeadas pela maioria das culturas e em que o grau de semelhança entre as cores é zero. Este conjunto de doze cores foi usado com sucesso em trabalhos anteriores [8] e é também utilizado na pesquisa de imagens de motores de busca, como no Google ou no Bing.4

As paletes resultantes com as cores mencionadas e mapeadas para as duas análises, cores dominantes e cores de pesquisa, podem ser consultadas na Figura 3.4 - Cores mencionadas e mapeadas para as duas questões respetivamente (estas paletes podem ser consultadas em ponto maior, em anexo).

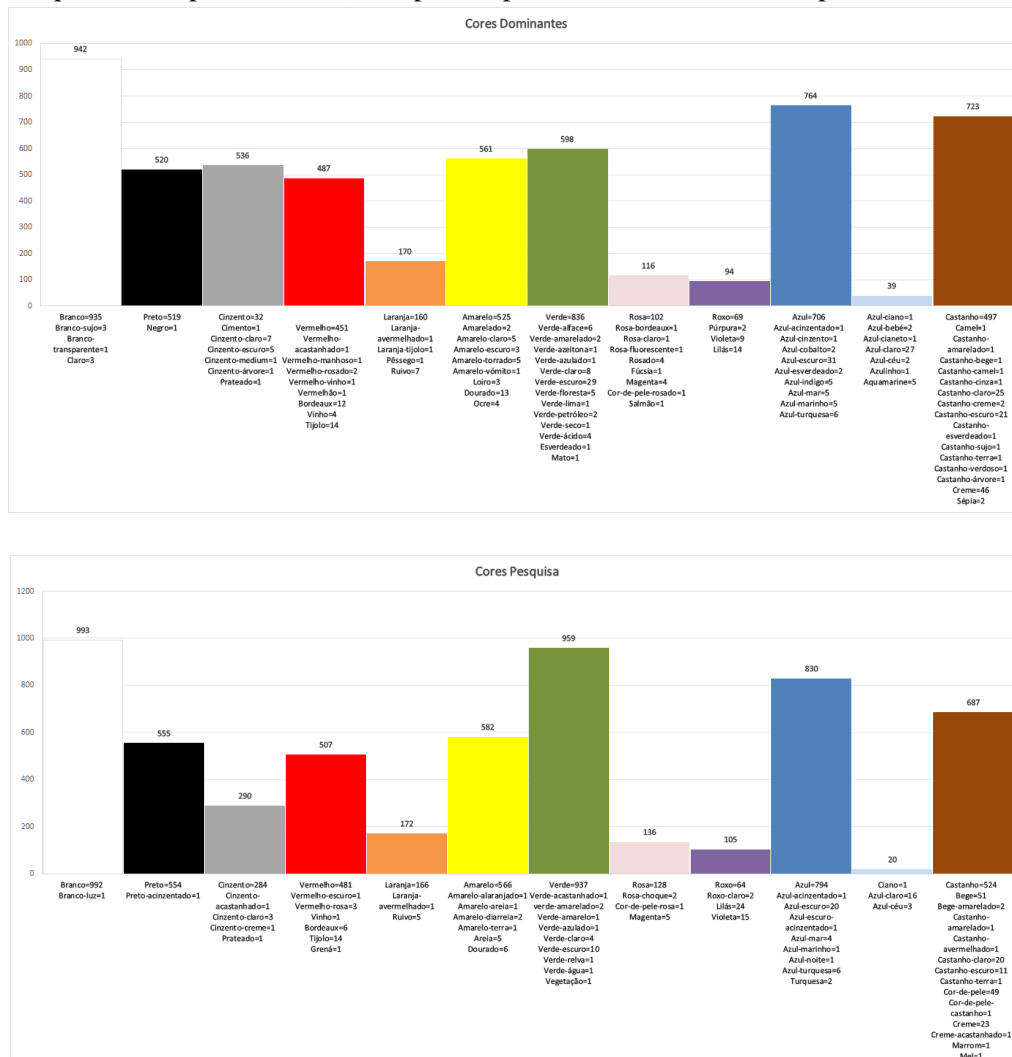


Figura 0.4 - Cores mencionadas e mapeadas para as duas questões.

Ao comparar os resultados de cores identificadas e mapeadas, que podem ser vistos na tabela de cores identificadas (ver Tabela 0.1) da análise de cores dominantes e da análise de cores de pesquisa, notamos que no geral (total de cores) as pessoas mencionaram mais cores na análise de pesquisa, que é refletido numa maior menção de cores e conversão para estas doze cores. As únicas cores da análise de pesquisa com valor menor que as das cores da análise dominante foram o Cinzento, Castanho e Ciano.

Tabela de Cores Identificadas			
Cor	Dominante	Pesquisa	Diferença
Branco	942	993	-51
Preto	520	555	-35
Cinzento	536	290	246
Vermelho	487	507	-20
Laranja	170	172	-2
Amarelo	561	582	-21
Verde	598	959	-370
Rosa	116	136	-20
Roxo	94	105	-11
Azul	764	830	-66
Ciano	39	20	19
Castanho	723	687	36
Total	5550	5836	-286

Tabela 0.1 - Tabela de Cores Identificadas nas Análises.

3.3.1. Concordância de Cores

Para cada imagem, identificamos as três cores mais mencionadas pelos utilizadores, para as duas questões. Nos casos de empate entre cores, consideramos cores adicionais, o que faz com que certas imagens tenham 4 ou 5 cores dominantes/pesquisa.

Depois de identificarmos as cores para cada imagem, verificamos qual a concordância entre as cores consideradas dominantes e as cores usadas para pesquisar com o objetivo de verificar se existia alguma diferença entre elas. Para isso, usamos a seguinte fórmula:

$$\frac{\#coresEmComum(Dominante, Pesquisa)}{\#MinCores(Dominante, Pesquisa)}$$

Ou seja, o número em comum de cores mais ditas que as duas questões tiveram a dividir pelo número mínimo de cores mais ditas entre as duas questões. Observamos que o denominador desta fórmula é sempre 3 para as 100 imagens. Assim sendo obtivemos os seguintes resultados.

Concordância (Dominante, Pesquisa)	#
1	70
0,67	28
0,33	2

Tabela 0.2 - Tabela de Concordância para cores mais identificadas entre questões.

Com estes resultados reparamos que em 70 imagens as respostas foram iguais, em 28 dos casos duas das cores apareciam nas mais mencionadas para as duas questões e em 2 casos apenas uma das cores coincidia entre si. Temos assim a taxa de concordância entre as questões:

$$\frac{(70 \times 1) + (0,67 \times 28) + (0,33 \times 2)}{100} = 89,33\%$$

Uma taxa de concordância entre as questões Q1 e Q2 de 89,33% representa um grau elevado de similaridade nas respostas. Concluimos assim que praticamente não há diferença para os utilizadores quando lhes é perguntado por cores dominantes de uma imagem e cores utilizadas para pesquisa da mesma. Ou seja, um possível algoritmo que sirva para identificar cores dominantes em imagens segundo a percepção humana, também servirá para evidenciar as cores que seriam utilizadas por um utilizador para pesquisar a mesma imagem.

3.4. Regiões de Foco

Nesta secção fazemos a análise das regiões para onde as pessoas olharam durante os testes, quer para a análise das cores dominantes quer para as cores de pesquisa. Foram criados heatmaps (mapas de regiões de foco), onde as zonas vermelhas representam as coordenadas mais olhadas pelos participantes.

3.4.1. Todas as Imagens

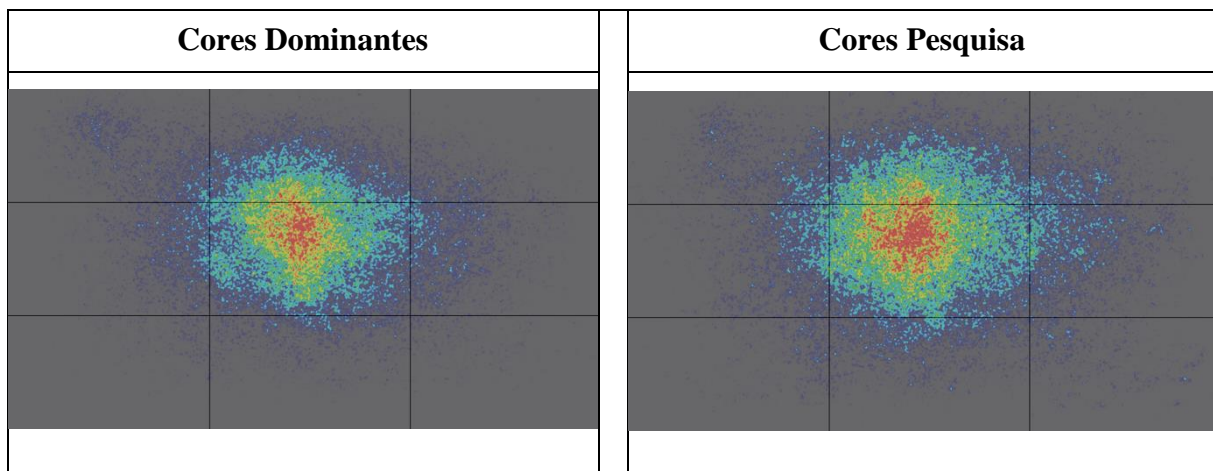


Tabela 0.3 - Heatmaps de todas as imagens para as duas análises.

Ao observar os dois heatmaps das duas análises, notou-se que os participantes focaram se maioritariamente no meio das imagens.

3.4.2. Categorias

Seguem se alguns exemplos de heatmaps de algumas categorias do conjunto.

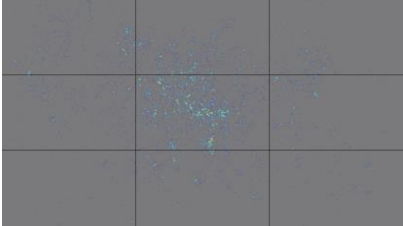
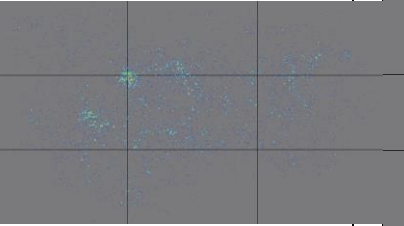
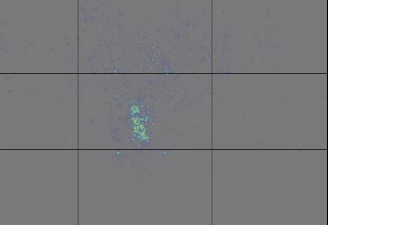
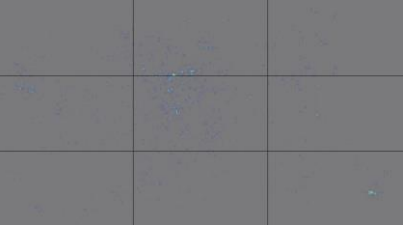
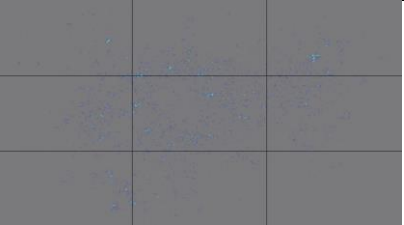
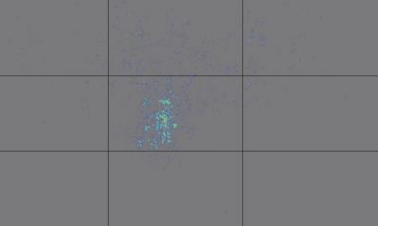
<i>Building</i>		<i>People</i>		<i>Snow</i>
Heatmaps de Cores Dominantes				
				
Heatmaps de Cores Pesquisa				
				

Tabela 0.4 - Heatmaps de algumas categorias para as duas análises.

3.4.3. Imagens Individuais

Seguem-se alguns exemplos de heatmaps de algumas imagens do conjunto.







Imagem #6	Imagem #58	Imagem #69
Heatmaps de Cores Dominantes		
		
Heatmaps de Cores Pesquisa		
		

Tabela 0.5 - Heatmaps de algumas imagens para as duas análises.

3.5. Discussão e Conclusões

Depois da análise dos *heatmaps* e das cores mencionadas conseguimos listar descobertas com o nosso estudo e conjunto de imagens que será importante na criação do algoritmo. As descobertas foram as seguintes:

1. Pessoas no geral, tendem a olhar mais para o centro das imagens.

Como pode ser visto nos dois *heatmaps* das imagens todas tanto para as cores dominantes como as cores de pesquisa, existe uma maior incidência no centro dos *heatmaps* (ver secção 3.4.1. Todas as Imagens.

2. Pessoas tendem a focar-se nas caras presentes nas imagens, mas as cores mencionadas são de elementos próximos das caras, mais concretamente cabelos e roupas.

Aqui podemos ver alguns *heatmaps* de imagens para categorias: Portrait, Girl e People, com resultados de análise de cores de pesquisa e as três cores mais mencionadas em cada uma.

Imagem #21	Imagem #82	Imagem #25	Imagem #72
Heatmaps de Cores Dominantes			
			
34% Preto	34% Castanho	34% Branco	34% Verde
20% Castanho	34% Verde	30% Preto	31% Azul
19% Verde	29% Branco	27% Castanho	17% Amarelo
Heatmaps de Cores Pesquisa			
			
32% Preto	33% Castanho	33% Branco	33% Verde
24% Verde	33% Branco	33% Preto	30% Azul
22% Rosa	32% Verde	16% Castanho	27% Amarelo

Tabela 0.6 - Heatmaps foco em imagens com caras.

Os heatmaps mostram um foco na cara das pessoas, mas as cores mais mencionadas para cada uma das análises são encontradas nos cabelos, roupas, fundo ou até nos lábios presentes nas imagens. O mesmo pode ser dito em relação à caras de animais e a cor do seu corpo.

3. Nas imagens só com natureza as coordenadas ficam mais dispersas. Não existe um objeto ou elemento de foco.

Como indicam os próximos heatmaps de exemplo, com as categorias sky, landscape e river por ordem, não existe um elemento claro onde as pessoas se possam focar.





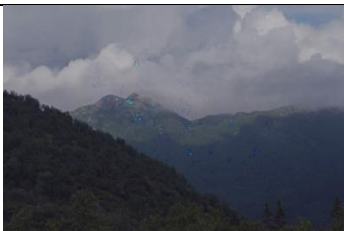

Imagem #1	Imagem #32	Imagem #99
Heatmaps de Cores Dominantes		
		
Heatmaps de Cores Pesquisa		
		

Tabela 0.7 - Heatmaps de natureza.

4. Pessoas tendem a olhar para focos de luz na imagem (sol, candeeiros, semáforos, etc.), mas as cores mais mencionadas nessas imagens poucas vezes representam o foco de luz.



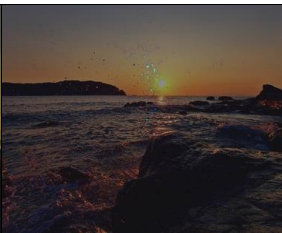
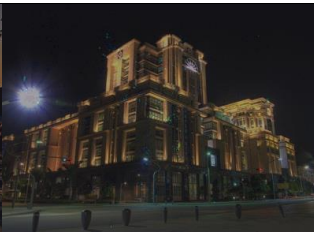




Imagem #18	Imagem #38	Imagem #77	Imagem #100
Heatmaps de Cores Dominantes			
			
37% Preto	39% Amarelo	33% Azul	32% Preto
28% Laranja	33% Preto	21% Amarelo	30% Amarelo
22% Amarelo	16% Branco	14% Laranja	13% Branco
Heatmaps de Cores Pesquisa			
			
32% Preto	33% Amarelo	28% Amarelo	23% Preto
27% Laranja	28% Branco	22% Azul	22% Amarelo
25% Amarelo	19% Preto	16% Laranja	18% Branco

Tabela 0.8 - Heatmaps presença de luz.

Nestes heatmaps pode-se verificar que os focos de luz são um elemento de foco mas nem sempre representam as cores mais ditas.

5. Pessoas tendem a olhar para o centro de uma flor, mas as cores mais mencionadas nessas imagens nem sempre representam o centro das flores.







Imagem #24	Imagem #26	Imagem #83
Heatmaps de Cores Dominantes		
		
36% Roxo	31% Vermelho	32% Vermelho
30% Verde	24% Rosa	29% Branco
16% Azul	22% Amarelo	21% Rosa
Heatmaps de Cores Pesquisa		
		
34% Roxo	28% Verde	24% Vermelho
33% Verde	24% Rosa	24% Branco
7% Azul	19% Vermelho	22% Rosa

Tabela 0.9 - Heatmaps de flores.

Nestes heatmaps da categoria Flowers conseguimos ver que as regiões de maior foco não coincidem com o centro de imagem mas que se aproximam com o centro das flores. Com as cores mais mencionadas para cada uma das análises vemos que apesar de o centro das flores ser uma região de foco nem sempre as cores presentes nessa zona fazem parte das cores mencionadas (ou seja mais percebidas).

6. Pessoas olham mais para a parte superior de edifícios.

Nos heatmaps exemplificativos seguintes, para as categorias Night e Architecture por ordem, podemos ver um maior foco nas regiões superiores dos edifícios presentes.



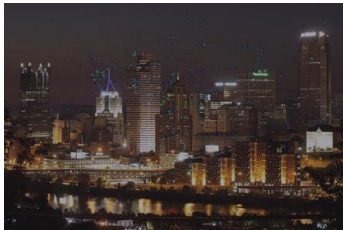

Imagem #12	Imagem #42
Heatmaps de Cores Dominantes	
	
Heatmaps de Cores Pesquisa	
	

Tabela 0.10 - Heatmaps edifícios.

Outras razões que podem ajudar a este efeito será a presença de luz nessas mesmas regiões (ver descoberta nº4), ou que a maioria das imagens presentes no dataset dá maior ênfase à parte superior dos edifícios.

7. Algumas pessoas leem o que está nas imagens.

Nestes heatmaps seguintes podemos ver algumas zonas de foco em letras, em placas, carros ou em edifícios. Quando participantes foram expostos aos seus resultados, apenas alguns manifestaram ter lido o que se encontrava nas imagens, outros com a pressão da tarefa imposta pelo estudo (encontrar três cores) não prestaram atenção a frases ou letras. Em certos casos, as letras são tão evidentes ou têm um tamanho que é difícil evitar ler.







Imagem #37	Imagem #91	Imagem #98
Heatmaps de Cores Dominantes		
		
Heatmaps de Cores Pesquisa		
		

Tabela 0.11 - Heatmaps de letras presentes.

8. Quanto mais pura for uma cor (sem adição de branco/preto) mais as pessoas se focam na cor.







Imagem #10	Imagem #35	Imagem #84
Heatmaps de Cores Dominantes		
		
34% Amarelo	28% Vermelho	35% Verde
22% Rosa	22% Castanho	33% Roxo
14% Laranja	16% Azul	13% Vermelho
Heatmaps de Cores Pesquisa		
		
24% Rosa	33% Vermelho	32% Verde
22% Amarelo	18% Azul	28% Roxo
14% Laranja/Castanho	15% Branco	22% Vermelho

Tabela 0.12 - Heatmaps de cores puras.

Com estes resultados, podemos verificar que algumas cores são escolhidas em detrimento de outras por serem puras. Na figura 10, as cores amarelo e rosa são escolhidas antes de quaisquer outras para a imagem em qualquer uma das análises. No caso da imagem 35 observamos a inclusão da cor azul nas três cores mais mencionadas, independentemente da análise, quando a cor azul representa uma parte mínima da imagem (apenas na bandeira). Analogamente para a imagem 84, a inclusão da cor vermelha (apenas presente nos sapatos) nas três mais mencionadas em detrimento de outras. Podemos observar outro exemplo similar com a imagem 77 (Ver descoberta nº4) em que a cor amarela é escolhida mais vezes que outras que podem parecer em maior quantidade.

9. Se a cor for quente mais as pessoas se focam nela.





Imagem #14		Imagem #60	
Heatmaps de Cores Dominantes			
			
29% Verde		28% Vermelho	
25% Vermelho		25% Branco	
25% Cinzento		18% Amarelo	17% Laranja
Heatmaps de Cores Pesquisa			
			
29% Cinzento		23% Vermelho	
29% Vermelho		22% Amarelo	
25% Verde		17% Laranja/Branco/Verde	

Tabela 0.13 - Heatmaps de cores quentes.

Na imagem 14 pode se verificar que a cor vermelha está entre a primeira ou segunda mais escolhidas para as duas análises, e essa cor só se encontra em um dos carros que representa uma pequena parte da imagem (carro este que se encontra perto do centro da imagem, ver descoberta nº1). Na imagem 60 não é de estranhar que a cor vermelha se apresente em primeiro lugar, mas sim que ambas as cores laranja e amarelo (cores com distinção mínima para esta imagem), sejam mencionadas tantas ou mais vezes que a cor verde ou outras.

Outros exemplos similares podem ser vistos em imagem de descobertas anteriores. Na descoberta nº2, imagem 72, a cor amarelo vem mencionada nas três mais mencionadas para ambas as análises e antes que a cor preto ou qualquer outra presente no fundo da imagem. Na descoberta nº 4, imagem 77, a cor amarela é vastamente mencionada e a cor laranja está presente nas três mais mencionada em vez de uma possível presença da cor preto. Na descoberta nº6, imagem 12, a cor amarelo é mais mencionada que outras cores como cinzento, castanho ou branco. Também na descoberta nº8, imagem 10, as cores amarelo e laranja estão presentes nas três mais mencionadas para ambas as análises.

Bailey et al. [12] observaram ainda que, segundo a percepção humana, objetos com cores quentes no geral têm maior probabilidade de parecer mais próximos na imagem que os restantes.

10. Pessoas dão mais importância às zonas focadas nas imagens.





Imagem #39	Imagem #80
Heatmap de Cores Dominantes	
	
Castanho 32%	Verde 36%
Branco 30%	Castanho 35%
Preto 15%	Azul 18%
Heatmap de Cores Pesquisa	
	
Branco 32%	Verde 34%
Castanho 32%	Castanho 26%
Verde 20%	Azul 24%

Tabela 0.14 - Heatmaps de cores em zonas focadas.

Na imagem 39, vemos que apesar da cor cinzento estar presente em grande parte e no fundo da imagem, não está presente nas três mais mencionadas em qualquer uma das análises. Acontece o mesmo para a imagem 80 com a cor branca, sendo esta preterida pela cor azul nas duas análises e visto que tem menor presença que a cor branca na imagem.

Estas descobertas são resumidas em algumas conclusões. As características de uma cor, tais como ser quente (visto em 9) ou ser cor pura (visto em 8) influenciam as respostas das pessoas. A localização da própria cor tem influência nas respostas obtidas, tais como a cor se encontrar no centro da imagem (visto em 1) ou se encontrar numa zona focada da imagem (visto em 10). É de salientar que, e como as descobertas 2,4 e 5 indicam, as zonas para as quais as pessoas olham nas imagens nem sempre representam as cores indicadas por estas. Estas últimas descobertas mais a descoberta 3, indicam que o conteúdo da imagem influencia o modo como as pessoas procuram por cores.

3.6. Síntese

Neste capítulo foi descrito como foi elaborado e conduzido o estudo, mais precisamente como foram reunidas as pessoas e imagens e o despiste de algumas destas duas, necessário para uma melhor recolha de informação. É também mencionado como são registadas as cores mencionadas pelas pessoas, como são capturadas as regiões para onde estas olham e ainda como foi calibrado e montado o eyetracker. Posteriormente é mostrada uma análise dos resultados dos estudos, que permitiu retirar conclusões para a criação do algoritmo proposto.

Capítulo 4.

Algoritmo de identificação de cores segundo a percepção humana

Neste capítulo descrevemos a proposta realizada para fazer a extração das cores dominantes numa imagem segundo a percepção humana. Após estudar o trabalho relacionado, e consultar os resultados do estudo com utilizadores conseguimos identificar passos importantes a incluir antes e durante o algoritmo, e que podem ser vistos nas secções que apresentamos a seguir.

4.1. Arquitetura

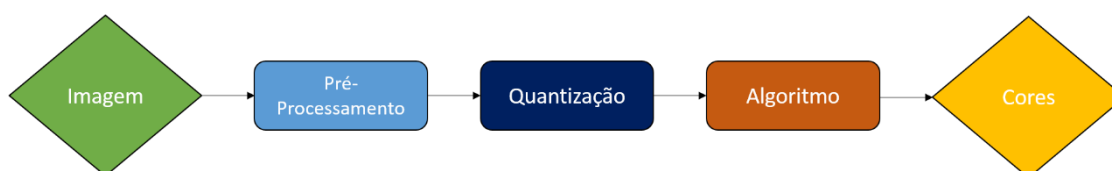


Figura 0.1- Arquitetura da Solução.

O algoritmo recebe de entrada qualquer imagem (fotografia ou não) e faz um pré-processamento desta que consiste em redimensionar e suavizar a imagem. Com a imagem pré-processada, o passo seguinte consiste em aplicar um algoritmo de quantização que propomos para converter cada cor dos pixéis na imagem em uma das 12 cores da paleta. Por fim o algoritmo faz a contabilização de cada cor na imagem, tendo em conta fatores descobertos durante o estudo com os utilizadores. São eleitas as 3 cores dominantes, ou mais em caso de empate.

4.2. Pré-processamento

Antes da identificação das cores dominantes segundo a percepção humana, é necessário realizar alguns ajustes às imagens, de modo a que o algoritmo as consiga correr de uma forma consistente e com o menor ruído possível nestas.

- **Redimensionar**

Para que possamos ter uma leitura semelhante de todas as imagens no algoritmo e para que este seja eficiente, as imagens devem ter tamanhos pequenos e semelhantes entre si. Para tal redimensionamos as imagens reduzindo para 400px a sua maior dimensão (comprimento ou largura) e redimensionamos em proporção a outra dimensão.

- **Suavizar**

Ao termos as imagens com o tamanho desejado é necessário reduzir o ruído existente na imagem que pode ter sido também introduzido pelo redimensionamento. Para tal aplicamos uma suavização gaussiana (Gaussian Smoothing ou Gaussian Blur). Este processo suaviza a imagem, tornando os contornos das “zonas” de cor mais suaves através da redução de detalhes.

4.3. Quantização para 12 cores

Para cada uma das doze cores presentes na paleta de Ware, foram criadas listas de cores com várias matizes, tonalidades e intensidades que pudessem ser traduzidas para uma das 12 cores representante. Cada instância de cor foi inicializada em valores RGB com ajuda de sites como <http://colorizer.org/> , <http://www.color-hex.com/> , <http://www.htmlcsscolor.com/> e <http://www.colorhexa.com/> que ajudaram na recolha dos valores para cada cor. Foram também feitas várias pesquisas de imagens de paletas de tonalidades e intensidades para cada uma das doze cores no Google para ajustar na criação destas listas.

Estas cores e as cores dos pixéis das imagens foram convertidas para o espaço CIE $L^*a^*b^*$, uma vez que, e como mencionado anteriormente, este espaço se aproxima da visão humana e é perceptualmente uniforme (ver Secção 2.2).

Para decidir qual a cor (das 12 da paleta) a atribuir a um pixel na imagem usa-se a menor distância das posições das cores instanciadas para com a posição da cor do pixel no espaço de cor CIE $L^*a^*b^*$. Se esta distância for menor que a de um *threshold* definido (5.0.) então o pixel tem essa cor representante, caso contrário é feita uma média para cada lista das doze cores representantes (resultando num ponto médio das cores no espaço CIE $L^*a^*b^*$) e a cor do pixel tomará a cor equivalente à cor representante com a distância menor (da média da lista de cores).

Na Figura 4.2 está uma representação simplificada de como é feita a quantização no nosso algoritmo. O ponto a rosa representa um pixel da imagem no espaço CIE $L^*a^*b^*$, os pontos a preto representam várias instâncias da cor verde (linhas pretas representam as distâncias entre as cores e o pixel), o ponto branco representa a média desses pontos (linha branca representa a distância entre a média da cor verde e o pixel).

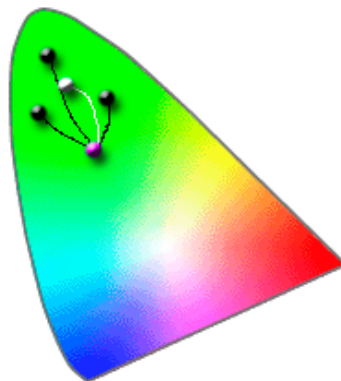


Figura 0.2- Espaço Cie $L^*a^*b^*$ em duas dimensões.

Resumindo a quantização:

1. Calcular cada distância (no espaço CIE $L^*a^*b^*$) da cor do pixel da imagem com cada cor presente em todas as 12 listas de cores. Escolher a menor distância.
2. Se essa distância for menor que o *threshold* pré-definido, então encontramos a cor do pixel. Termina a procura da cor do pixel.
3. Caso contrário, calculam-se as distâncias da cor do pixel com a média de cada uma das 12 listas de cores. A menor das 12 distâncias representa a cor do pixel. Termina a procura da cor do pixel.

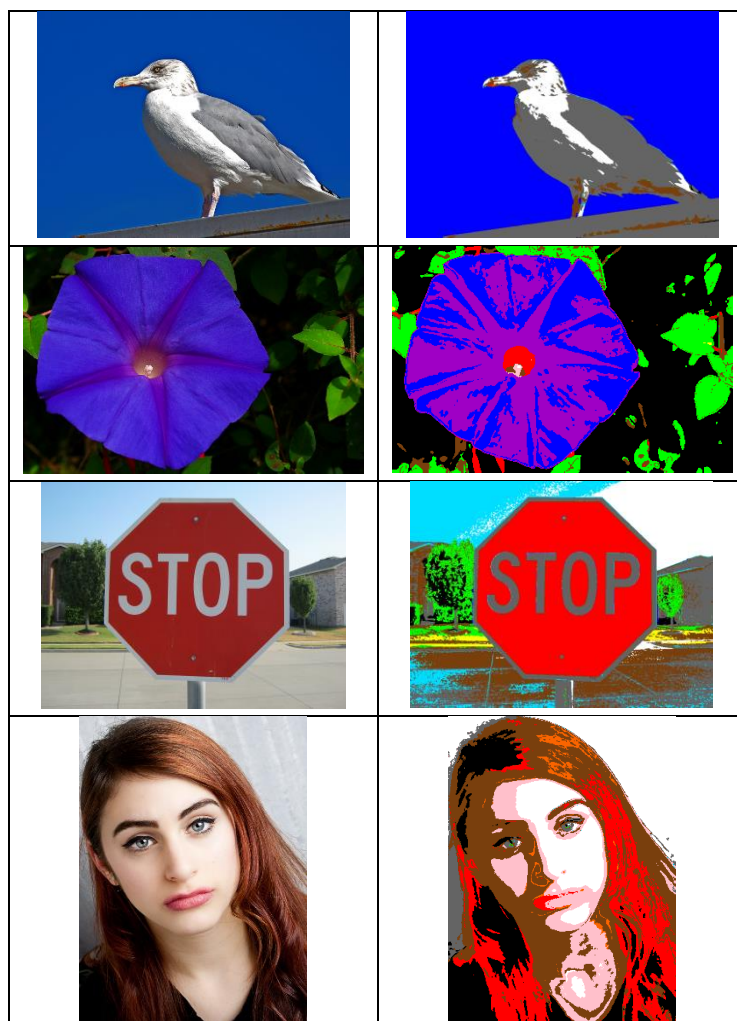


Figura 0.3 - Exemplo de imagens do conjunto submetidas à quantização para as 12 cores.

A Figura 0.3 apresenta exemplos de imagens que foram quantizadas com o nosso método. As primeiras duas revelam resultados positivos, ao passo que as últimas duas mostram resultados menos desejados (a imagem do sinal tem contornos e letras cinzentas quando se queria branco, enquanto que na imagem da cara as cores do cabelo e da pele deveriam variar menos), estes resultados podem dever-se à nossa percepção inicial errada da cor dos píxeis nas imagens ou ao algoritmo proposto da quantização.

4.4. Algoritmos

Nesta secção são apresentados os diferentes algoritmos que concebemos e como estes fazem a extração de cores dominantes nas imagens.

4.4.1. Algoritmo base

Este algoritmo, ao qual chamamos algoritmo base, consiste numa contagem simples das cores pixel por pixel na imagem. No final as cores são ordenadas de forma decrescente, elegemos as três cores mais dominantes e retorna a percentagem destas nos píxeis da imagem. É de realçar que este algoritmo não tem qualquer tipo de percepção humana associada.

4.4.2. Algoritmo base com votos

Este algoritmo, ao qual chamamos algoritmo base com votos, consiste na divisão da imagem numa matriz de 16 blocos (4x4), onde cada bloco pode contribuir com um voto para cada uma das suas três cores dominantes (menos, caso não haja três cores e mais, em caso de cores empatadas). Em essência é como se cada bloco executasse o algoritmo base (sem votos) elegendo cada um as suas três cores dominantes. Se o número de pixéis de uma cor num dado bloco corresponder a um valor menor que 10% dos pixéis do bloco, então essa cor é descartada e a cor não pode receber votos. No final avaliam-se e somam-se os votos dos vários blocos, elegendo-se como as três cores dominantes na imagem as que obtiverem mais votos. Este algoritmo, tal como o anterior, também não têm qualquer tipo de percepção humana associada.

4.4.3. Fatores para a percepção humana

Para incluir a percepção humana nos dois algoritmos atrás mencionados, baseamo-nos em algumas das conclusões do estudo do Capítulo 3. Assim aplicamos alguns fatores às cores dos pixéis, sejam estes aplicados a nível do pixel ou a nível do bloco. Os fatores identificados e utilizados a partir das descobertas do estudo descrito no Capítulo 3 são: cores quentes, cores puras, centro da imagem e focagem.

1) A nível do pixel:

Estes fatores são independentes da localização do pixel na imagem.

- **Cores Puras**

Como resultado do estudo, foi verificado que as pessoas olham e mencionam em primeiro lugar as cores puras. Sendo assim, sempre que é identificado um pixel de cor pura é dado mais peso a esse pixel e por conseguinte à cor representante na contagem total das cores. Uma cor pura é uma cor que tem pouca adição da cor branca e da cor preta, valores estes que são representados no espaço HSV pelas variáveis Saturation e Value (ver secção 2.2.2).

Para identificar se uma cor é pura, foi convertida a cor do pixel para o espaço HSV, e verificamos se os valores de S e V eram maiores ou iguais a 75².

- **Cores Quentes**

Outro resultado do estudo verificado foi o de que as pessoas olham e mencionam as cores quentes. Então foi dado mais peso aos pixéis de cores quentes e por consequente à cor representante na contagem total das cores. Para a cor de um pixel ser considerada quente esta tem que ser da cor Vermelho, Laranja ou Amarelo e com Saturation e Value superiores a 25³.

2) A nível do bloco:

Estes fatores estão diretamente relacionados com a localização do pixel nos blocos da imagem.

- **Centro da Imagem**

Outro resultado do estudo indicou que as pessoas olham principalmente para o centro

^{2 3} Resultado obtido de forma empírica

das imagens. Assim sendo, damos mais peso aos pixéis que estão nos quatro blocos centrais da imagem.

- **Zonas Focadas**

Como visto também nos resultados do estudo as pessoas olham principalmente para as zonas focadas das imagens. Para verificarmos se a imagem está focado aplicamos aos 16 blocos da imagem um filtro Laplaciano, que faz a detecção de contornos nas imagens. Quantos mais contornos tem um bloco mais focado este está. Assim sendo, foi dado mais peso aos pixéis pertencentes aos blocos focados da imagem.

4.4.4. Algoritmo de Percepção Humana

Este terceiro algoritmo, ao qual chamamos de Percepção Humana, funciona de modo semelhante ao algoritmo base (ver secção 4.4.1), ou seja faz contagem das cores dos píxeis na imagem, ordena e devolve as percentagens das cores mas com os pesos atribuídos aos píxeis provenientes dos fatores vistos na secção 4.4.3.

4.4.5. Algoritmo de Percepção Humana com votos

Este foi o quarto e último algoritmo que criamos funciona de modo semelhante ao algoritmo base com votos (ver secção 4.4.2), ou seja divide a imagem em 16 blocos e cada um elege as suas 3 cores dominantes para uma soma final dos votos, mas com os pesos devidos atribuídos aos píxeis provenientes dos fatores vistos na secção 4.4.3.

4.5. Síntese

Neste capítulo apresentamos a arquitetura da solução e os seus módulos: Pré-processamento necessário da imagem, a quantização desta para 12 cores e os algoritmos que desenvolvemos para testar. Concebemos algoritmos com e sem percepção humana, e para cada um destes uma versão que conta simplesmente as cores dos píxeis em toda a imagem e outra que faz a contagem das cores por votos de blocos na imagem resultantes da divisão da imagem em uma matriz de 16 blocos. São ainda indicados os fatores a incluir nos algoritmos de percepção humana.

Capítulo 5.

Avaliação Experimental

Neste capítulo fazemos a avaliação dos algoritmos e fatores a utilizar propostos para a extração de cores dominantes nas imagens.

5.1. Procedimento Experimental

Para testar cada algoritmo usamos o dataset de 100 imagens descrito no 0. Qualquer um destes algoritmos retorna as 3 cores com mais peso para cada imagem. No final comparamos os resultados de cada algoritmo com as cores mais mencionadas pelos utilizadores no estudo com os utilizadores. O objetivo é ter um resultado próximo das 300 cores dominantes (100 imagens x 3 cores dominantes).

5.2. Resultados dos Algoritmos Base

Ao utilizarmos os algoritmos base (sem percepção humana) para as 100 imagens do dataset, foram obtidos os seguintes resultados:

	Sem Voting		Com Voting	
A. Cores Dominantes	162	54%	174	58%
A. Cores de Pesquisa	159	53%	174	58%

Tabela 0.1 - Tabela com resultados de algoritmos base.

Como podemos ver pela tabela acima, o algoritmo simples sem voting e sem percepção humana consegue um resultado de 54% na análise das cores dominantes e 53% na análise das cores de pesquisa, para as 100 imagens do nosso dataset, valores bastante próximos entre si. Há no entanto uma subida de 4% e 5% respetivamente, quando usamos o algoritmo com votos (ainda sem adição de fatores que vão contribuir com a percepção humana).

5.3. Resultados de cada fator

De modo a obter resultados nos dois algoritmos que fazem uso da percepção humana mais próximos com os resultados do estudo do 0 foi necessário fazer uma análise aos fatores a utilizar. Aplicamos valores iguais de uma escala entre 0 e 50, a todos os fatores e corremos os algoritmos à vez para retirar os resultados.

5.3.1. Análise do fator de Cores Quentes

Segue-se a análise do fator de cores quentes com as cores dominantes resultantes do estudo com utilizadores. Para esta tabela e para as tabelas seguintes, a linha a cinzento representa os valores base e os valores a negrito representam os resultados máximos obtidos.

Fator Quente	Sem Voting	Com Voting
0	162	174
1	159	171
2	161	163
3	163	164
5	166	171
10	167	172
15	161	173
20	161	173
25	159	173
30	156	172
35	154	174
40	155	175
45	154	173
50	152	173

Tabela 0.2 - Tabela de análise ao fator de cores quentes com cores dominantes.

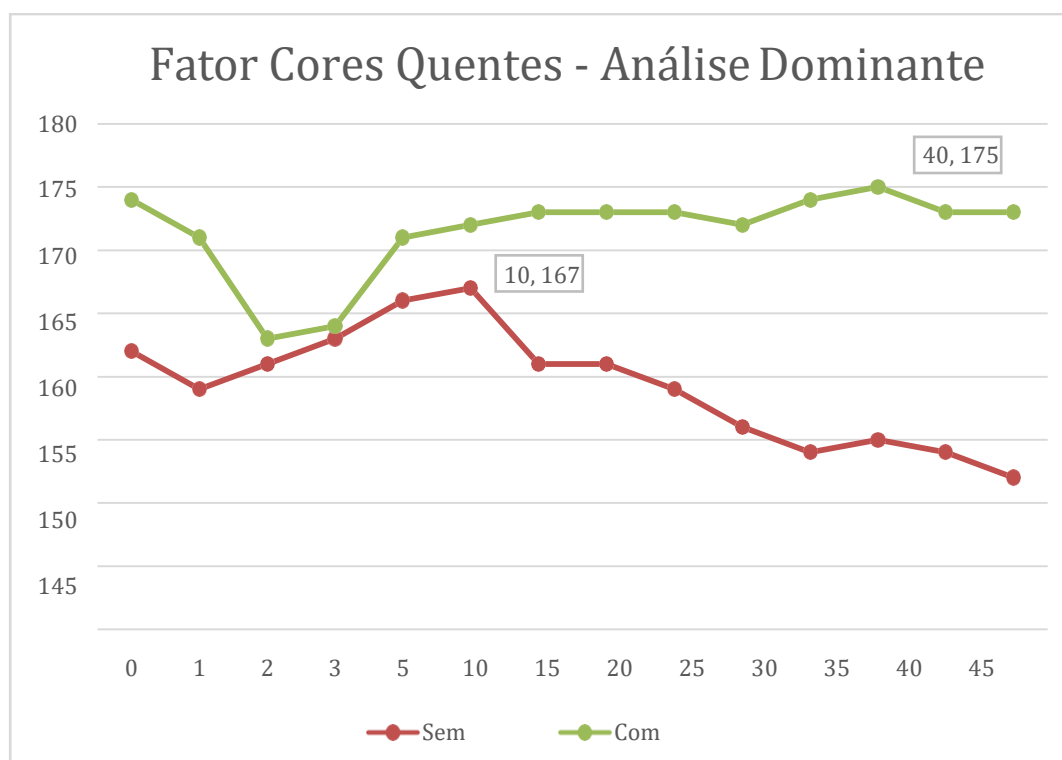


Figura 0.1 - Gráfico da análise do fator de cores quentes com cores dominantes.

Aqui mostramos a análise do fator de cores quentes com as cores de pesquisa resultantes do estudo com utilizadores.

Fator Quente	Sem Voting	Com Voting
0	159	174
1	157	172
2	158	163
3	159	163
5	161	170
10	164	171
15	162	172
20	161	172
25	159	173
30	158	172
35	156	174
40	157	175
45	156	174
50	154	174

Tabela 0.3 - Tabela de análise ao fator de cores quentes com cores de pesquisa

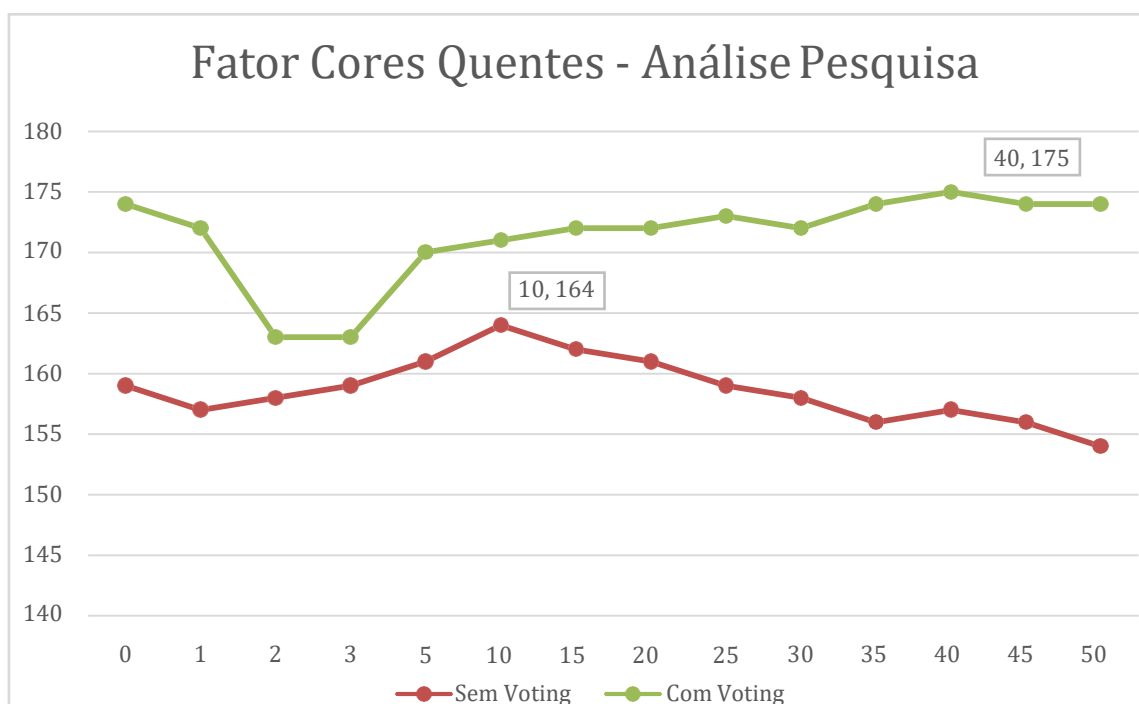


Figura 0.2 - Gráfico de análise do fator de cores quentes com cores de pesquisa.

O fator quente faz variar os resultados obtidos e apesar de os resultados nem sempre serem melhores que nos algoritmos bases, para as duas análises e para os dois algoritmos o resultado máximo é obtido quando usamos um valor diferente de zero (10 no caso dos algoritmos sem voting e 40 para os algoritmos com voting).

5.3.2. Análise do fator de Cores Puras

Mostramos aqui a análise do fator de cores puras com as cores dominantes resultantes do estudo com utilizadores.

Fator Pura	Sem Voting	Com Voting
0	162	174
1	163	174
2	163	175
3	163	176
5	163	177
10	165	176
15	168	176
20	169	172
25	171	174
30	170	174
35	169	175
40	169	176
45	169	176
50	169	176

Tabela 0.4 - Tabela de análise ao fator de cores puras com cores dominantes.

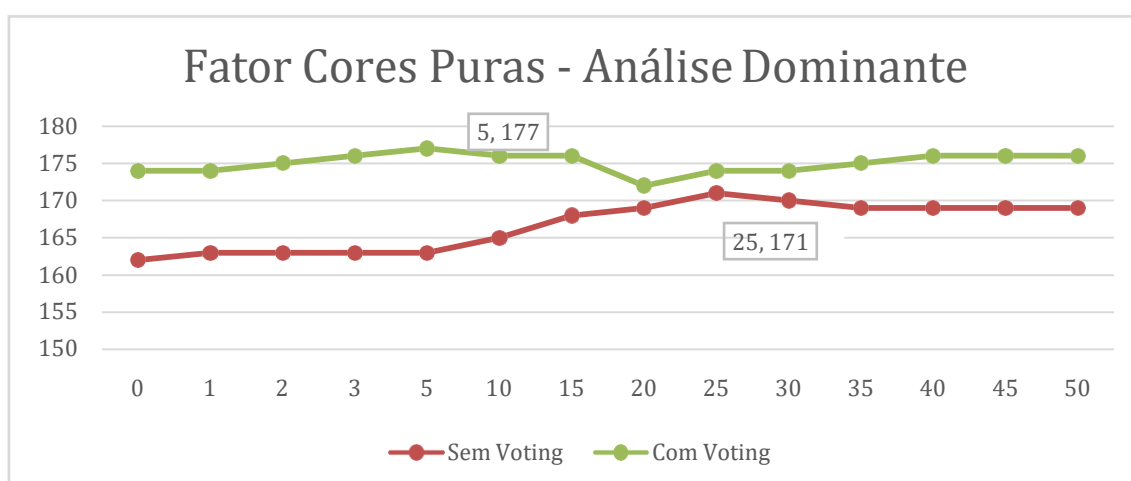


Figura 0.3 - Gráfico de análise do fator de cores puras com cores dominantes.

Segue-se a análise do fator de cores puras com as cores de pesquisa resultantes do estudo com utilizadores.

Fator Pura	Sem Voting	Com Voting
0	159	174
1	159	175
2	159	176
3	159	177
5	161	178
10	163	178
15	166	179
20	166	175
25	168	176
30	167	176
35	166	177
40	166	178
45	166	177
50	166	177

Tabela 0.5 - Tabela de análise ao fator de cores puras com cores de pesquisa.

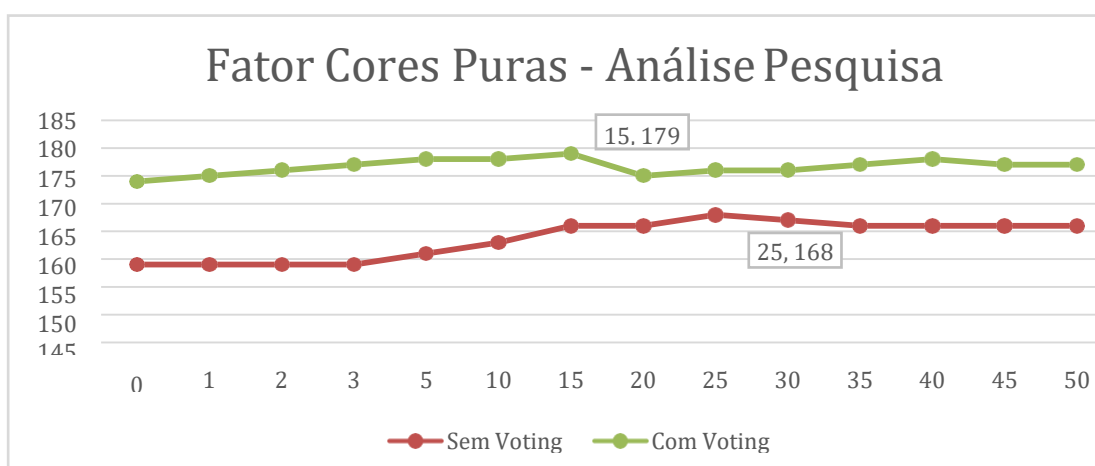


Figura 0.4 - Gráfico de análise do fator de cores puras com cores de pesquisa.

O fator de cores puras faz variar os resultados obtidos e independentemente do valor a utilizar os resultados obtidos são sempre melhores que os resultados dos algoritmos base. Para o algoritmo sem voting o resultado máximo obtido é quando usamos o valor 25, já para os algoritmos com voting existe um máximo quando é usado o valor 5 na análise de cores dominantes e um máximo quando é usado o valor 15 na análise das cores de pesquisa. Assim sendo vamos atribuir uma média destes valores (10) para usar como resultado máximo do fator de cores puras do algoritmo com voting no estudo seguinte.

5.3.3. Análise do fator de Centro da Imagem

Mostramos aqui a análise do fator de cores no centro das imagens com as cores dominantes resultantes do estudo com utilizadores.

Fator Meio	Sem Voting	Com Voting
0	162	174
1	161	169
2	158	168
3	159	168
5	159	169
10	161	169
15	161	170
20	162	170
25	163	170
30	163	170
35	164	170
40	166	170
45	164	170
50	163	170

Tabela 0.6 - Tabela de análise ao fator de cores no centro da imagem com cores dominantes.

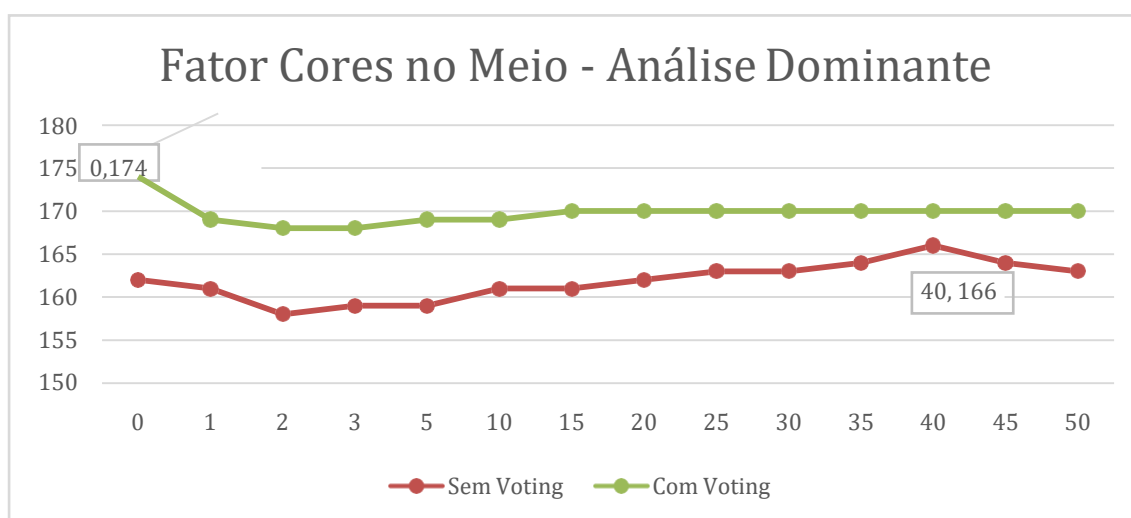


Figura 0.5 - Gráfico de análise do fator de cores no centro da imagem com cores dominantes.

Segue-se a análise do fator de cores no centro das imagens com as cores de pesquisa resultantes do estudo com utilizadores.

Fator Meio	Sem Voting	Com Voting
0	159	174
1	156	170
2	158	169
3	160	169
5	159	169
10	161	169
15	162	170
20	163	170
25	164	170
30	164	170
35	165	170
40	167	170
45	164	170
50	163	170

Tabela 0.7 - Tabela de análise ao fator de cores no centro da imagem com cores de pesquisa.

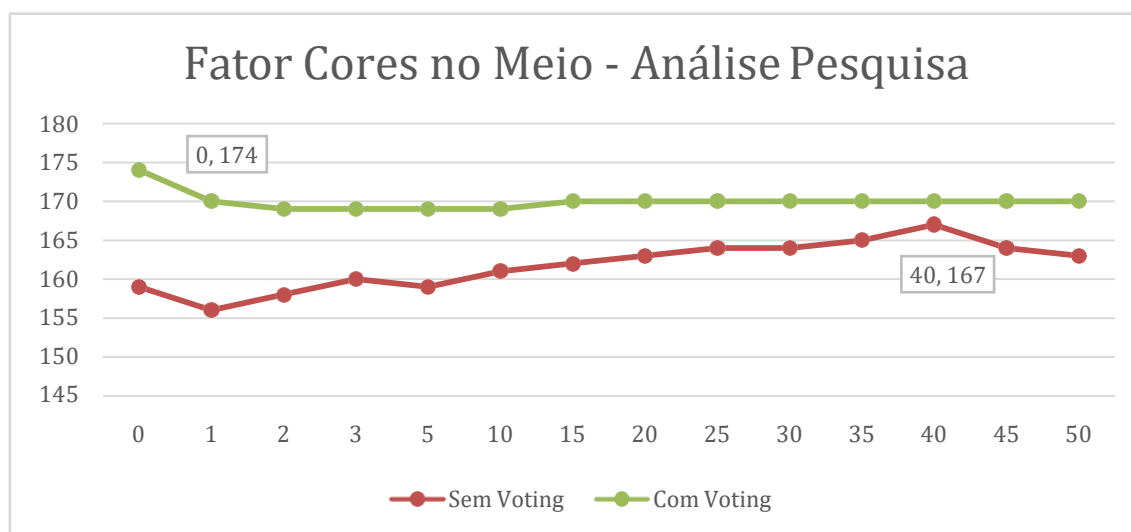


Figura 0.6 - Gráfico de análise do fator de cores no centro da imagem com cores de pesquisa.

O fator de cores situadas no meio da imagem faz variar os resultados obtidos e apesar de os resultados nem sempre serem melhores que no algoritmo base para o algoritmo sem voting, estes são sempre piores para o algoritmo com voting, nas duas análises. Assim sendo o máximo obtido para o fator de cores situadas no meio da imagem foi de 40 para o algoritmo sem voting e 0 para o algoritmo com voting, nas duas análises.

5.3.4. Análise do fator de Zonas Focadas

Fazemos aqui a análise do fator de cores em zonas focadas das imagens com as cores dominantes resultantes do estudo com utilizadores.

Fator Foque	Sem Voting	Com Voting
0	162	174
1	162	170
2	165	168
3	165	168
5	165	168
10	159	168
15	159	167
20	159	168
25	159	168
30	159	168
35	158	168
40	159	168
45	160	168
50	159	167

Tabela 0.8 - Tabela de análise ao fator de cores em zonas focadas com cores dominantes.

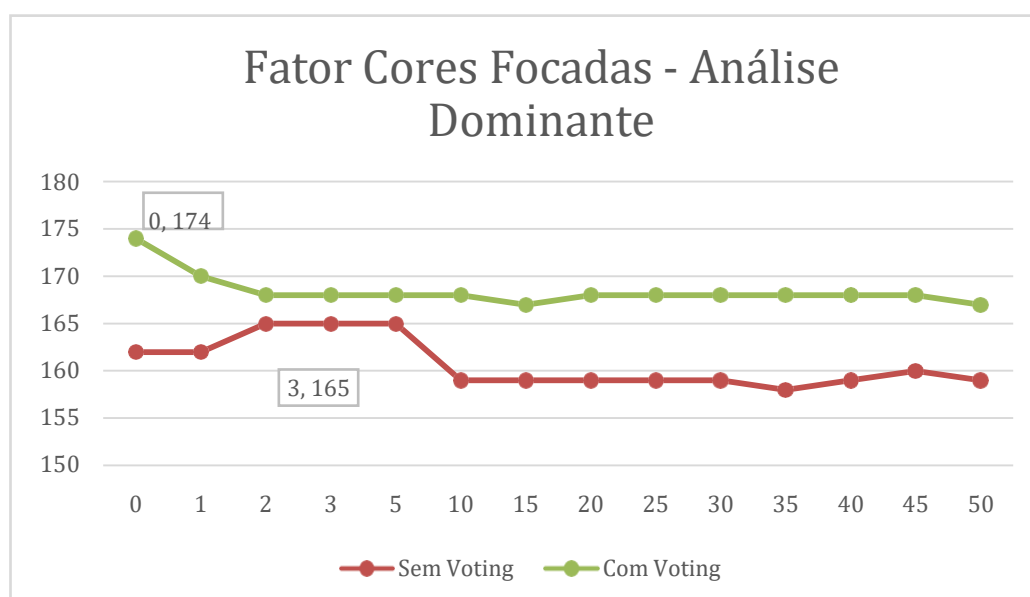


Figura 0.7 - Gráfico de análise do fator de cores em zonas focadas na imagem com cores dominantes.

Segue se a análise do fator de cores em zonas focadas nas imagens com as cores de pesquisa resultantes do estudo com utilizadores.

Fator Foque	Sem Voting	Com Voting
0	159	174
1	160	169
2	163	167
3	163	167
5	163	167
10	156	167
15	157	167
20	156	168
25	156	167
30	156	167
35	155	167
40	156	167
45	157	167
50	156	166

Tabela 0.9 - Tabela de análise ao fator de cores em zonas focadas com cores de pesquisa.

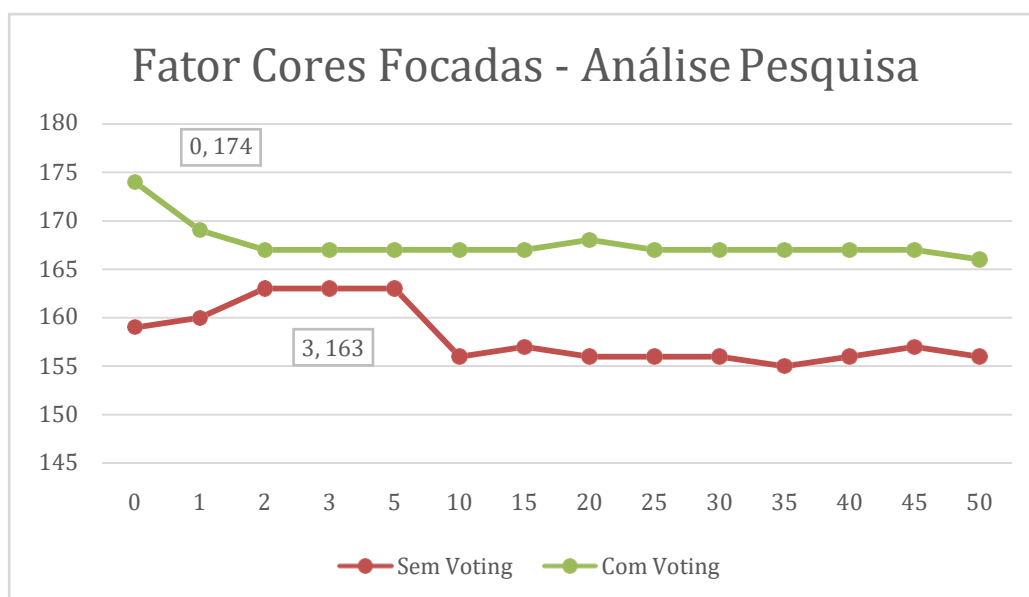


Figura 0.8 - Gráfico de análise do fator de cores em zonas focadas na imagem com cores de pesquisa.

O fator de cores situadas em zonas focadas da imagem faz variar os resultados obtidos e apesar de os resultados nem sempre serem melhores que no algoritmo base para o algoritmo sem voting, estes são sempre piores para o algoritmo com voting, nas duas análises. Assim sendo o máximo obtido para o fator de cores situadas em zonas focadas da imagem foi de 3 para o algoritmo sem voting e 0 para o algoritmo com voting, nas duas análises.

5.4. Resultados dos algoritmos de percepção humana

Reunimos os máximos para cada fator das duas análises (cores dominantes e cores de pesquisa), de modo a realizarmos um estudo combinatório de máximos e verificar se podemos melhorar os resultados já obtidos dos algoritmos. Para facilitar a apresentação do estudo e os seus resultados vamos representar os fatores por letras (Q: fator de cores quentes, P: fator de cores puras, M: fator de cores no meio da imagem, F: fator de cores em zonas focadas na imagem).

Os máximos de cada fator para as duas análises respetivamente foram:

Cores Dominantes			
Sem voting		Com voting	
Q	10	Q	40
P	25	P	5
M	40	M	0
F	3	F	0

Tabela 0.10 - Tabela de valores máximos para fatores com cores dominantes.

Cores de Pesquisa			
Sem voting		Com voting	
Q	10	Q	40
P	25	P	15
M	40	M	0
F	3	F	0

Tabela 0.11 - Tabela de valores máximos para fatores com cores de pesquisa.

Reparamos que os valores para as duas análises são similares com a exceção do fator das cores puras.

Para tentar obter um valor que excedesse os já obtidos, fizemos um estudo mais aprofundado com as diferentes combinações de máximos dos fatores:

QPMF	QPM	QP PM MF
	QPF	QC PF
	QMF	QF
	BMF	

Tabela 0.12 - Tabela de configurações de máximos a experimentar.

Os resultados das configurações para o algoritmo sem voting, e ordenados por percentagem de acerto, foram os seguintes:

Configuração	Valores Utilizados	Resultados AD	Percentagem de Acerto	Resultados AP	Percentagem de Acerto
P	<0,25,0,0>	171	57%	168	56%
PF	<0,25,0,3>	169	56,3%	167	55,7%
QPF	<10,25,0,3>	169	56,3%	163	54,3%
QPM	<10,25,0,3>	169	56,3%	163	54,3%
PM	<0,25,40,0>	168	56%	167	55,7%
QM	<10,0,40,0>	168	56%	166	55,3%
QP	<10,25,0,0>	168	56%	165	55%
Q	<10,0,0,0>	167	55%	164	54,7%
M	<0,0,40,0>	166	55,3%	167	55,7%
QPMF	<10,25,40,3>	165	55%	162	54%
F	<0,0,0,3>	165	55%	163	54,3%
QF	<10,0,0,3>	165	55%	160	53,3%
QMF	<10, 0, 40, 3>	164	54,7%	161	53,7%
PMF	<0, 25,40,3>	163	54,3%	161	53,7%
MF	<0, 0, 40,3>	159	53%	159	53%

Tabela 0.13 - Configurações de máximos testadas no algoritmo sem voting.

A melhor configuração para o algoritmo sem voting é quando apenas tem em conta o fator das cores puras, resultando em uma taxa de acerto de 57% para a análise das cores dominantes e em 56% para análise das cores de pesquisa. Apenas a última configuração (MF) obtêm valores piores do que o algoritmo base sem voting.

Relembramos a concordância de cores entre análise de cores dominantes e cores de pesquisa (em 3.3.1) para considerar como 10 o valor do fator de cores puras para análise de combinações de máximos no algoritmo com voting (média entre as duas análises) e que neste algoritmo só dois dos fatores tiveram máximos para valores diferentes de zero.

Configuração	Valores Utilizados	Resultados AD	Percentagem de Acerto	Resultados AP	Percentagem de Acerto
P	<0,10,0,0>	176	58,7%	178	59,3%
QP	<40,10,0,0>	176	58,7%	176	58,7%
Q	<40,0,0,0>	175	58,3%	175	58,3%

Tabela 0.14 - Configurações de máximos testadas no algoritmo com voting.

A melhor configuração para o algoritmo com voting é também quando apenas tem em conta o fator das cores puras, resultando em uma taxa de acerto de 58,7% para a análise das cores dominantes e em 59,3% para análise das cores de pesquisa. Todas as configurações obtiveram melhores resultados que o algoritmo base com voting.

5.5. Análise e Discussão

Nos algoritmos base, ou seja sem percepção humana o número de imagens acertadas foi:

	Sem Voting		Com Voting	
A. Cores Dominantes	162	54%	174	58%
A. Cores de Pesquisa	159	53%	174	58%

Tabela 0.15 - Cores acertadas pelos algoritmos base.

Com o estudo aos algoritmos de percepção humana que propomos os melhores resultados obtidos foram:

	Sem Voting		Com Voting	
A. Cores Dominantes	171	57%	176	58,7%
A. Cores de Pesquisa	168	56%	178	59,3%

Tabela 0.16 - Cores acertadas pelos algoritmos de percepção humana.

Notou-se uma subida de 3% para os resultados dos algoritmos sem voting e uma subida menos significativa para os algoritmos com voting.

Uma possível razão pela qual o algoritmo com voting não precisa da maioria dos fatores e tem uma subida mínima para a sua versão com percepção humana, é a deste algoritmo ser mais suscetível a devolver cores empatadas do que na percentagem de cores resultante do algoritmo sem voting. Assim, mais cores são comparadas com as que foram retiradas do estudo com as pessoas

No geral quase todas as combinações de algoritmos de percepção humana com algum dos fatores associado resulta num melhor resultado que os dos algoritmos bases, no entanto recomendamos utilizar o algoritmo com voting com o fator de cores puras a 10.

5.6. Síntese

Neste capítulo fizemos uma avaliação aos algoritmos propostos no capítulo anterior. Para tal começamos por explicar em detalhe como é feita esta avaliação. De seguida fizemos uma avaliação individual e combinatória dos fatores para os algoritmos de percepção humana. Por fim comparamos os resultados e mencionamos o algoritmo e a configuração que ofereceu melhores resultados.

Capítulo 6.

Conclusões e Trabalho Futuro

Neste capítulo apresentamos sucintamente todo o trabalho efetuado durante a dissertação e alguns meios para melhorar o algoritmo proposto.

6.1. Sumário da dissertação

Com o intuito de desenvolver um algoritmo que identificasse as cores dominantes numa imagem do ponto de vista da percepção humana começamos por estudar alguns trabalhos já existentes que lidam com cores e fazem a identificação das cores dominantes em imagens (0), onde retiramos algumas técnicas que viriam a ser úteis. Além disso quisemos perceber como é feita a percepção da cor pelos humanos, o que são os modelos de cor e quais os que nos trariam mais vantagens para o objetivo da dissertação.

Após obter a informação sobre as técnicas, precisamos de recolher os nossos dados através de um estudo (0) realizado com utilizadores sobre o modo como estes identificam as cores dominantes em imagens, quais as cores que utilizariam para encontrar uma imagem e quais as zonas da imagem para onde olham durante a escolha das cores. Para esse efeito criamos um conjunto de 100 imagens seguindo alguns critérios e utilizamos esse conjunto numa aplicação que desenvolvemos, que com ajuda do dispositivo eyetracker que rastreia o movimento dos olhos, permitiu retirar conclusões para a criação do algoritmo proposto. Além de perceber que cores dominantes as pessoas identificam nas imagens, conseguimos identificar os nomes das cores mencionadas para criar uma paleta reduzida de cores, as zonas que as pessoas olharam enquanto identificaram as cores dominantes e reparamos em algumas das características das cores e das suas localizações, bem como os motivos que levaram as pessoas a olhar mais para elas.

Já com as técnicas e dados recolhidos, planificamos e formulamos a arquitetura da solução (0). Mostramos quais os passos importantes a incluir antes e durante o algoritmo. Oferecemos uma proposta de quantização de cores e quatro algoritmos para extração de cores dominantes em imagens, com e sem percepção humana de modo a compará-los.

Explicamos e fizemos então a avaliação dos algoritmos e dos fatores a utilizar que oferecem a percepção humana (0). Comparamos os resultados e sugerimos o melhor algoritmo e a sua melhor configuração de fatores a utilizar na extração de cores dominantes (com base na percepção humana).

6.2. Trabalho Futuro

Algumas técnicas utilizadas no nosso algoritmo podem ser melhoradas, tais como:

- Detecção de Zonas Desfocadas/Focadas

No nosso algoritmo dividimos a imagem numa matriz 4x4 (16 blocos) para fazer a deteção de blocos mais e menos focados, o que representa uma forma pouco precisa de evidenciar as zonas mais focadas de uma imagem. Uma melhor forma de o fazer é usar mapas de saliência que devolvem as posições exatas dos pixéis nas zonas focadas na imagem.

- **Lista de Cores**

Existe um número vasto de cores possíveis que podem ser instanciadas e fazerem parte da lista de cores do algoritmo para devolver uma melhor quantização das cores numa imagem. Esta lista pode ser atualizada de forma a conter mais informação e devolver cores com mais precisão para cada pixel.

Algumas das 12 cores utilizadas têm mais peso que outras (utilizadores raramente dizem ciano para a cor ciano nas imagens, optam mais pela cor azul; limites de cores como as cores quentes ou cinzento com branco/preto também geram dúvidas).

Os pesos que damos às cores nas imagens dependem do seu conteúdo. No estudo que fizemos com utilizadores detetamos por exemplo que os utilizadores olham mais para o centro das flores ou para as roupas e cabelos das pessoas. Existem ferramentas que nos permitem detetar o conteúdo das imagens, tais como:

- **Clarifai**

É uma API que nos devolve uma lista das categorias prováveis de uma imagem.

- **Microsoft Vision e Face**

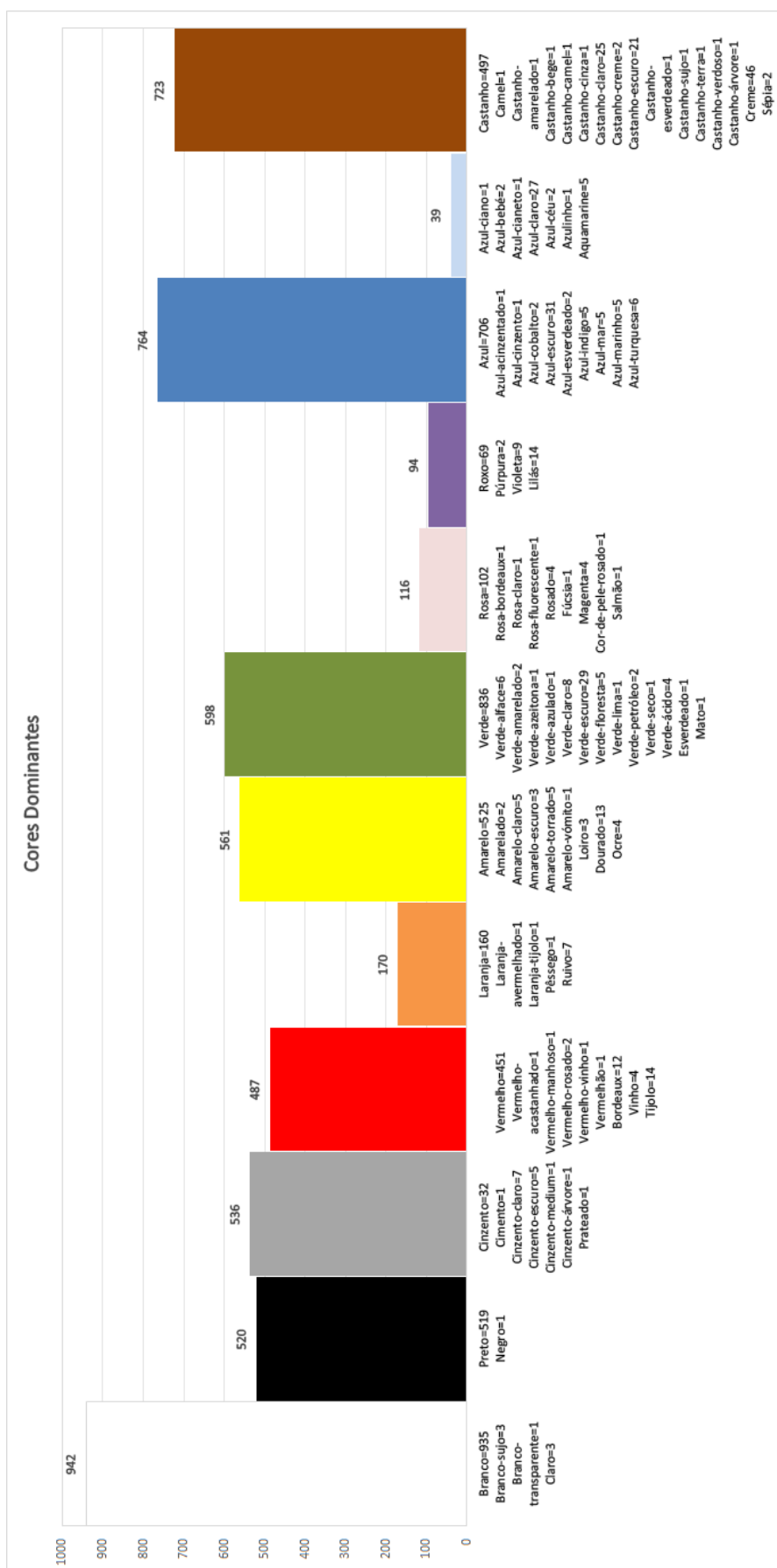
A API da Microsoft além de devolver uma lista com categorias prováveis de uma imagem (em conjunto com a sua probabilidade), mostra ainda posições e atributos relevantes das caras das pessoas nas imagens, como o tamanho, posição e cor das íris dos olhos; a cor do cabelo ou a localização da própria cara em si.

Com esta informação podemos adicionar novos fatores ao nosso algoritmo ou dar pesos diferentes aos fatores já existentes em função do tipo de elementos na imagem.

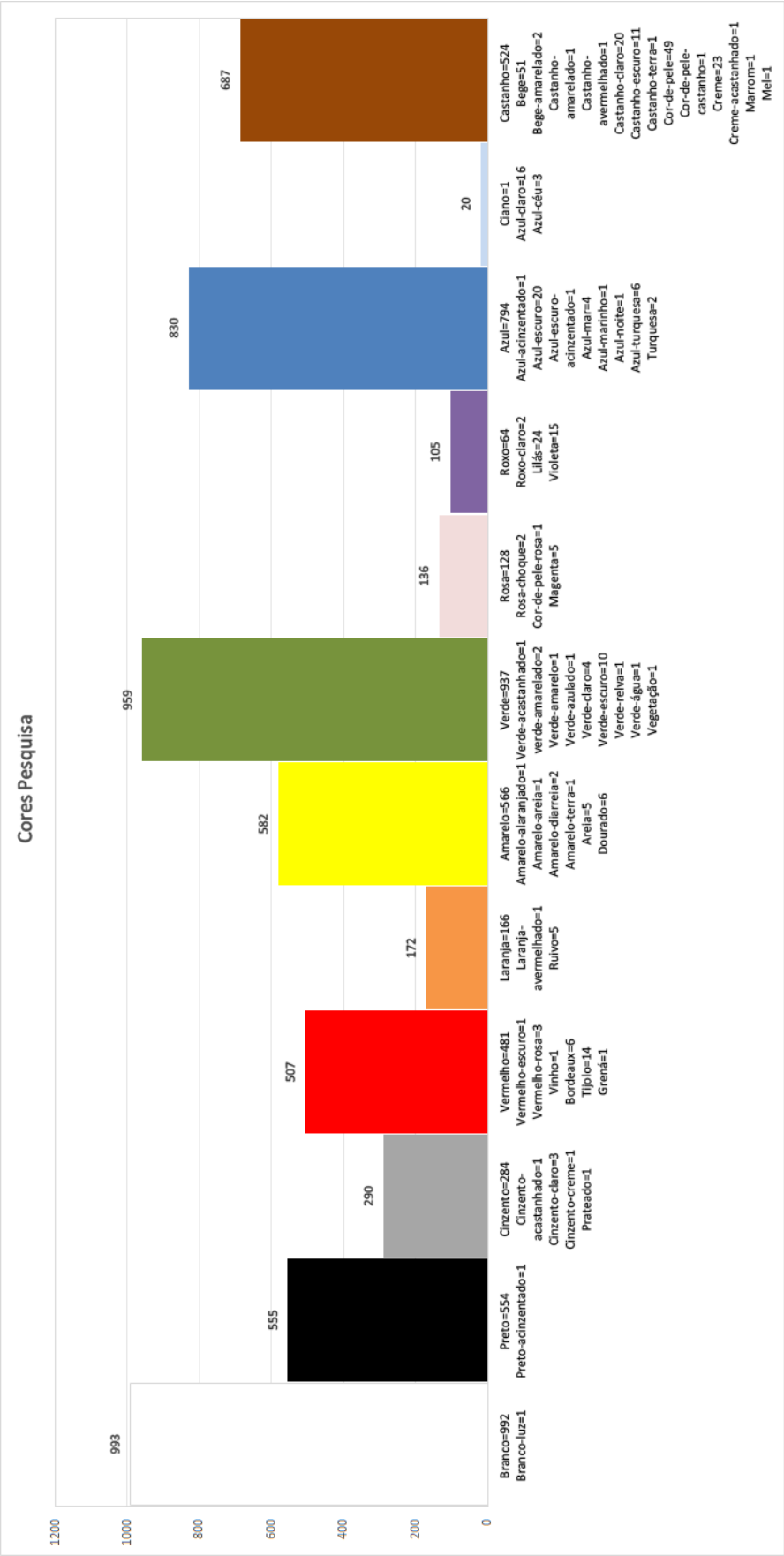
Bibliografia

- [1] Ahmed Talib, Massudi Mahmuddin, Husniza Husni, Loay E, George “A weighted dominant color descriptor for content-based image retrieval”, J. Vis. Comun. Image Represent, pp. 345-360, 2013.
- [2] Yang Lifang, Huang Xianglin, Lv Rui, Lv Hui “An Effective Similarity Measurement Algorithm for Dominant Color Feature Matching in Image Retrieval”, Applied Mechanics and Materials, Vols. 182-183 , pp. 1169-1173, 2012.
- [3] Shih-Fu Chang, Thomas Sikora, Atul Puri, “Overview of the MPEG-7 standard”, IEEE Trans. Cir. and Sys. for Video Technol. 11, pp. 688-695, 2001.
- [4] Peng Wang, Dongqing Zhang, Gang Zeng, Jingdong Wang, “Contextual Dominant Color Name Extraction for Web Image Search”, Proceedings of the 2012 IEEE International Conference on Multimedia and Expo Workshops (ICMEW '12), pp. 319-324, 2012.
- [5] Jen-Rainer Ohm, Leszek Cieplinski, Heon Jun Kim, Santhana Krishnamachari, B. S. Manjunath, Dean S. Messing, Akio Yamada, “The MPEG-7 Visual Description Framework - Concepts, Accuracy, and Applications,” CAIP, pp. 2-10, 2001.
- [6] Noor A. Ibraheem, Mokhtar M. Hasan, Rafiqul Z, Khan, Pramod K. Mishra, “Understanding Color Models: A Review”, ARPN Journal of Science and Technology, 2012.
- [7] LeGrand H. Hardy, Gertrude Rand, M. Catherine Rittler, "Tests for the detection and analysis of color-blindness.", Journal of The Optical Society of America, Vol. 35, pp. 268–271, 1945.
- [8] José C. Amante, Manuel J. Fonseca, “Fuzzy Color Space Segmentation to Identify the Same Dominant Colors as Users”, 18th International Conference on Distributed Multimedia Systems (DMS'12), 2012.
- [9] Edward Y. Chang, Beita Li, Chen Li, “Toward perception-based image retrieval”, Content-based Access of Image and Video Libraries Proceedings. IEEE Workshop on Content-based Access of Image and Video Libraries, pp. 101-105, 2000.
- [10] Colin Ware, “Information Visualization: Perception for Design.”, Morgan Kaufmann Publishers Inc., 2004.
- [11] Giuliana Ramella, Gabriella Sanniti di Baja, “Color histogram-based image segmentation”, Proceedings of the 14th international conference on Computer analysis of images and patterns - Volume Part I (CAIP'11), pp. 76-83, 2011.
- [12] Reynold Bailey, Cindy Grimm, Christopher Davoli, "The Effect of Warm and Cool Object Colors on Depth Ordering", Tech. Rep. 17, Washington University in St. Louis, 2006.

Anexos



Anexo 1 - Paleta Resultante do Estudo das Cores Dominantes.



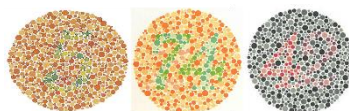
Anexo 2 - Paleta Resultante do Estudo das Cores de Pesquisa.

Guião para estudo com utilizadores

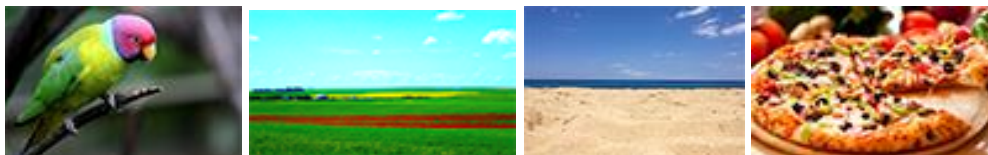
Os seguintes passos devem ser tomados para a realização de uma sessão de estudo com sucesso.

1. Mostrar apresentação com os testes de Ishihara para detetar possível daltonismo do utilizador.

Perguntar que números o utilizador consegue identificar nas três imagens (as respostas devem ser: 5; 74; 42; respetivamente).



2. Caso o utilizador tenha errado uma destas o teste termina aqui.
3. Registrar as seguintes informações sobre o utilizador:
 - Idade: _____
 - Sexo: F ☐ M ☐
 - Portador de
 - Lentes de contacto: S/N
 - Óculos: S/N
4. Ajustar e calibrar o Eyetracker com o utilizador.
5. Lançar a aplicação para recolha dos movimentos do olhar usando o Eyetracker em modo de apresentação (usando todo o ecrã se possível) com as cem imagens que testam a percepção humana das cores.



6. Numa folha (mais se preciso) e para cada imagem na apresentação:
Perguntar até três cores dominantes, por ordem, que o utilizador encontra na imagem a ser mostrada e registar todas as cores referidas pelo utilizador.
7. Registrar notas adicionais que ache fundamentais sobre a experiência decorrida.

Anexo 3 - Guião para estudo com utilizadores.